

University of St Andrews



Full metadata for this thesis is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

**The influence of social structure
on vocal signatures in group-living
resident killer whales (*Orcinus orca*)**

by

Anna E Nousek

**A thesis submitted in fulfillment of
the requirements for the degree of Master of
Philosophy**

**School of Biology
University of St Andrews**

submitted 29 September 2004



TH E 937

ABSTRACT

Constant association with a specific set of individuals appears to be advantageous for some species. This lifestyle then generates many modifications on individual behavior, especially to coordinate relationships between individuals within the group, and vocal signals are likely to be shaped by social interactions as they may play a role in mediating social roles. Northern resident killer whales (*Orcinus orca*) use group-specific signatures within their stable social system but to date no study has thoroughly examined calling behavior amongst individuals within different family groups. Thus, the potential for individual-specific signatures has not been identified. This study aimed to describe the effects of their rigid social structure on the vocal behavior of individuals through comparing the intra-individual call similarity with their inter-individual similarity. A towed beamforming array of hydrophones allowed calls to be ascribed to specific individuals. This study quantified call similarity through the ability of a pattern recognition neural network to discriminate between frequency contours of individuals within the same matrilineal group, in different matrilineal groups but the same vocal clan and across different clans. Calls were found to contain features distinct to an individual, but these features are less prominent than group-specific features. Similarity of both the high and low frequency contours was strongly affected by the social relationship of pairs of whales, with the network having the highest error rate for same whale comparisons, a lower error rate for comparisons within a matrilineal group, and the lowest error rate for comparisons

between individuals in different matrilineal groups. The correlation between social relationship and acoustic similarity suggests social learning as the mechanism for transmission and that group identity has a stronger influence on the time-frequency structure of stereotyped killer whale calls than individual identity. The social bonds between groups of individuals appear to take precedence over individual relationships within this population.

Declaration

I, Anna Nousek, hereby certify that this thesis, which is approximately 30000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date 06/05/05 signature of candidate

I was admitted as a research student in October 2003 and as a candidate for the degree of Master of Philosophy in October 2003; the higher study for which this is a record was carried out in the University of St Andrews between 2003 and 2004.

date 06/05/05 signature of candidate

I hereby certify that the candidate has fulfilled the condition of the Resolution and Regulations appropriate for the degree of Master of Philosophy in the University of St Andrews and the candidate is qualified to submit this thesis in application for that degree.

date 6/5/5 signature of advisor

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

date 06/05/05 signature of candidate

TABLE OF CONTENTS

Abstract	i
Declaration	iii
Table of Contents	iv
Acknowledgements	vii
 <i>Chapter 1 – General Introduction</i>	
Social structure in cetaceans	1-1
Resident killer whales	1-1
Transient killer whales	1-3
Sperm whales	1-4
Bottlenose dolphins	1-5
Relationship between social structure and vocal behavior	1-6
Vocal trends in individual-based societies	1-7
Vocal trends in group-based societies	1-8
Vocalizations in cetacean societies	1-10
Vocal patterns in bottlenose dolphins	1-10
Vocal patterns in sperm whales	1-11
Vocal patterns in transient killer whales	1-12
Vocal patterns in resident killer whales	1-13
 <i>Chapter 2 – Neural networks for pattern recognition used to discriminate between calls of resident killer whales with different levels of social association</i>	
Data collection	2-1
Recording individuals in the field	2-1
Ascribing calls to an individual	2-2
Time-frequency contours	2-3
Contour extraction	2-4
Measurement of call similarity	2-6
Neural network pattern recognition	2-6
Network structure	2-6
Neural network justification	2-8
Comparison bootstrapping	2-9
Selection criteria for network input	2-10
Input size trials	2-11
Randomization trials	2-12
Conclusion	2-14

Chapter 3 – The investigation of optimal network selection to assess the relationship between vocal signatures and resident killer whale social structure

Introduction.....	3-1
Methods	3-3
Results.....	3-8
Hand selection method	3-10
A pod.....	3-10
W pod.....	3-12
Summary	3-13
Random selection method.....	3-14
A pod.....	3-14
W pod.....	3-17
Summary	3-17
Comparison of selection methods.....	3-19
Discussion	3-20

Chapter 4 – The evaluation of temporal influence on acoustic similarity between individuals with different social relationships

Introduction.....	4-1
Methods	4-5
Results.....	4-7
Random selection method	4-7
Selection from different group bouts	4-8
Comparison of selection methods.....	4-11
Discussion	4-12

Chapter 5 -Northern resident vocal similarity quantified through comparisons of separate components within and between call types

Introduction.....	5-1
Methods	5-5
Results.....	5-7
Within Call Type Analysis.....	5-8
Low Frequency Component.....	5-8
A pod.....	5-8
W pod.....	5-8
Summary	5-8
High Frequency Component.....	5-9
A pod.....	5-9
W pod.....	5-11
Summary	5-11
Within Call Type Summary.....	5-13

Chapter 5 – Continued

Within Shared Repertoire Analysis	5-13
Low Frequency Component.....	5-13
A pod	5-14
W pod.....	5-17
Summary	5-17
High Frequency Component.....	5-18
A pod	5-19
W pod.....	5-22
Summary	5-22
Within Shared Repertoire Summary.....	5-23
Unshared Repertoire Analysis.....	5-24
Low Frequency Component.....	5-24
A & W pods	5-24
High Frequency Component.....	5-25
A & W pods	5-25
Unshared Repertoire Summary	5-27
Comparison of Repertoire Analyses	5-28
Discussion	5-29
Patterns within call type.....	5-29
Patterns within shared repertoire	5-30
Patterns within unshared repertoires	5-32
Patterns across repertoires	5-33
Signature system of resident killer whales	5-33
Social learning.....	5-36
Important considerations	5-38
Conclusion	5-40

Chapter 6 – General Discussion

Study Goals.....	6-1
Signature systems.....	6-3
Group-specific.....	6-3
Individual-specific.....	6-5
Future work.....	6-7
Spectrographic analysis.....	6-7
Perceptual studies in the field.....	6-10
Social structure in other cetaceans.....	6-11
Conclusion	6-12

References	ix
-------------------------	-----------

Appendix I: Spectrograms by Individual.....	xxiv
--	-------------

Appendix II: Summary of Available Field Recordings	xli
---	------------

ACKNOWLEDGEMENTS

This project has involved many people up to this point, and I would like to thank all of those who have contributed along the way. First of all, I wish to give extensive thanks to my supervisor Dr. Patrick Miller who not only gave me an incredible opportunity, which allowed me to enjoy a different and exciting place, but also provided excellent advice and instruction at this starting point in my career. I would like to thank all of those who were involved with recording, processing and cataloguing these calls, of which Ari Shapiro and Andrea Hodgins-Davis deserve special mention because they are the only two whom I have had the pleasure of meeting and/or working with, and all of those who contributed when this project was involved with Dr. Miller's degree. My involvement in the project began at the Spoken Language System Department of the Massachusetts Institute of Technology, and I would like to thank the following individuals there. Dr. Chao Wang deserves considerable credit for developing the contour extraction program and solving all of the difficulties that arose from using their computing power from the other side of the Atlantic, and Dr. Stephanie Seneff for serving as my official advisor while I was working there.

I also wish to thank Professor Peter Slater for supervising me while Patrick was traveling to work on several different projects, giving wonderful advice and enhancing my understanding of not only animal behavior but also British culture. Although I only met him briefly, Dr. Volker Deecke helped me considerably by developing the code to run the neural networks and having the creativity to apply this technique to killer whale vocal behavior. I would also like

to thank the usual suspects that attend Peter's lunchtime meetings for both their insights on current scientific literature and their sparkling mealtime conversation; Sally Ward, Henrik Brumm, Lorraine Marshall-Ball and Luke Rendell. I appreciate the time and comments of both of my examiners, Vincent Janik and A. Rus Hoelzel, who provided a fresh perspective on this work in its written version. I would also like to specifically thank Hugo Rainey for company and crumpets during the long, cold and dark evenings and weekends that we both spent in the lab in order to finish and Nicola Quick for the occasion field experience as well as supplying information about events at the Sea Mammal Research Unit.

My final thanks goes to all of those who supported me emotionally along the way: Ross McGregor, my wonderful parents, and my two brilliant sisters Lisa and Katrina. Your efforts are not taken for granted, and I consider myself extremely fortunate to receive the love, encouragement and support of such wonderful people.

CHAPTER 1:

General Introduction

Social Structure in Cetaceans

In many species, living in a group has been shown to have benefits for the individual, such as increasing predator defense, foraging success, and mating success, as well as reducing energetic demands (Krause & Ruxton, 2002). If these groups are to be maintained, individuals must interact to coordinate their behavior. The relationship between these interactions, or the social structure, can take on many different forms based on the needs of individuals within their environment. Cetaceans present an excellent example of how diverse these structures can be even between species with similar ecological constraints. Within each cetacean society, sociality influences individual behavior and leads to the evolution of varied methods for the transfer of information. This chapter will review social structures in killer whales and compare them with those of two other well-studied species.

Resident killer whales

Inhabiting the waters off British Columbia, resident killer whales (*Orcinus orca*) have a clearly grouped social structure, which has been investigated through the use of extensive photo identification studies (Whitehead, 1995). These whales are highly recognizable using visual identification of dorsal fin shape, saddle patch variations and other natural markings and each animal has been labeled alphanumerically to create a catalogue of the entire population (Bigg, 1982;

Balcomb *et al.*, 1982). Studies following the movements of individuals over a long-term period have demonstrated the consistent distribution of individuals and that there are few changes in association patterns between them (Ford *et al.*, 2000). Through these studies, a study of acoustic similarity and genetic comparisons, the genealogy of this population of whales has been developed into the following levels. Individuals live in small groups described as matrilineal units, formed from three to ten whales all related to the eldest female 'matriach' of the group. Photo-identification data has shown that both male and female individuals travel with their natal matrilineal unit throughout their lifetime (Heimlich-Boran, 1986b; Olesiuk *et al.*, 1990; Ford *et al.*, 2000) although the exact extent of matrilineal unit fidelity is not known as genetic data suggests some male dispersal (Hoelzel *et al.*, 2002). Several matrilineal units with related ancestry and more infrequent associations form the next level of social interaction, known as the pod. Within the photo-identification database, individuals of different pods are named with different letters (Ford *et al.*, 2000). Groups of pods that associate with each other, albeit less frequently than individuals within a pod, are said to belong to the same clan. Separate clans include individuals of different ancestral lines of killer whales, although low genetic variation throughout the resident population suggests a fairly recent matrilineal relationship (Hoelzel & Dover, 1991). Finally, the killer whales off the western coast of North America can be divided into three communities defined by their geographic location that generally do not overlap (Ford *et al.*, 2000). The 'southern' community inhabits the waters of southern Vancouver

Island and Puget Sound, while the 'northern' community ranges from the northern area off Vancouver Island to southern Alaska. The third community inhabits the waters off the southern coast of Alaska. Later studies of genetic similarity have demonstrated that individuals at more distant levels of social structure are less closely related to each other (Barrett-Lennard, 2000; Hoelzel *et al.*, 1998; Hoelzel & Dover, 1991). These whales have been shown to feed only on fish, primarily salmon, and follow the seasonal salmon migrations through British Columbian waters (Nichol & Shackleton, 1996; Heimlich-Boran, 1986a). Foraging in these large groups facilitates locating patchy resources (Krause & Ruxton, 2002), and the abundance of salmon can support the large number of these predators within this small region (Ford, 1991; Baird & Whitehead, 2000; Connor *et al.*, 2000a). Similar resident populations are associated with the salmon populations off southern Alaska (Saulitis *et al.*, 2000) and the herring populations near Norway (Domenici *et al.*, 2000).

Transient killer whales

Social structure can even vary within a species, as the presence of a sympatric population of killer whales in these same waters demonstrates. Transient killer whales also inhabit the waters off western North America but feed on marine mammals (Ford *et al.*, 1998), a feeding specialization that could be facilitated by a very different type of social structure (Connor *et al.*, 2000a; Baird & Dill, 1996). Seals and small cetaceans are another food source in the Pacific Northwest, but these prey items cannot support a large number of individuals

foraging in the same area. A study of transient social structure showed that bonds between individuals remain strong, but all individuals cannot remain within their natal group due to the high energetic cost of this type of foraging. Therefore further generations will disperse to ensure that groups remain small (Baird & Whitehead, 2000; Baird & Dill, 1996). This extreme specialization in two very different prey types could be due to the lack of a single effective hunting technique for both marine mammals and fish (Barrett-Lennard *et al.*, 1996). In any case, it may have also contributed to a second type of social structure and consequent marked genetic differences (Hoelzel & Dover, 1991).

Sperm Whales

As with the social structure of transient killer whales, sperm whales (*Physeter macrocephalus*) demonstrate different levels of social fidelity, potentially influenced by feeding requirements. Female whales appear to remain at lower latitudes and travel in social groups that consist of a number of individuals and their offspring, with male offspring dispersing at maturity (Christal & Whitehead, 2001). Adult male sperm whales may travel to more productive higher latitudes which better satisfy their higher energy demands (Whitehead, 2003). This social structure may also yield other advantages, as the presence of multiple individuals may allow female sperm whales to better care for calves as well as defend themselves against predators. Sperm whales dive to depths of nearly a kilometer in order to locate prey, and young whales likely do not have the ability to dive alongside their mothers. Instead of leaving it alone at

the surface, other females of the group will accompany a calf while its mother is foraging (Arnborn & Whitehead, 1989; Whitehead, 2003). Grouping may also reduce the danger of predation as sperm whales have been observed clustering together in defensive formation to deter attacking predators (Pitman *et al.*, 2001; Arnborn *et al.*, 1987). Male sperm whales have a much less well-documented social structure although the stranding of multiple male sperm whales seems to indicate that social bonds exist, although at a different level to those within female groups (Whitehead, 2003).

Bottlenose dolphins

Bottlenose dolphins (*Tursiops truncatus*) have completely different type of social structure, one that is described as a 'fission-fusion' society. In this type of society, animals preferentially associate with specific individuals, but these bonds develop and change frequently (Connor *et al.*, 2000b). Individuals have been shown to maintain strong social bonds between mothers and calves and between pairs of males (Connor *et al.*, 2001; Sayigh *et al.*, 1998) but other interactions are apparently less stable. Individuals disperse at a young age, and in some populations will only associate for intermittent periods with one or two specific individuals during adulthood (Connor *et al.*, 1992; Connor *et al.*, 2000a; Connor *et al.*, 2001; Smolker *et al.*, 1992). It appears that there is no advantage to permanent groups, although the extent of individual association varies between dolphin populations and further study will be necessary to eliminate the

possibility of a group-based social structure within some populations of bottlenose dolphins (Connor *et al.*, 2000b).

These three species demonstrate three levels of social interactions within the same ecological constraints of the ocean and, as they are the best studied of cetacean species, demonstrate the best example for each type of social structure. Social structures have been reported for several other species of odontocete but they have not been studied as extensively (Amos *et al.*, 1993). For example, beluga or white whales, *Delphinapterus leucas*, appear to live in large groups in coastal areas following large fish stocks (de March & Postma, 2003; Heide-Jorgensen *et al.*, 2003). Studies of pilot whales show that neither sex disperses, so again there appears to be a benefit to a stable social structure (Amos *et al.*, 1993). Each species has evolved in order to maximize their chance of survival within the marine environment.

Relationship between Social Structure and Vocal Behavior

Vocalizations serve to convey information amongst individuals when other senses are of limited use, particularly where visibility is low, chemical transmission is limited or individuals are separated over large distances (Hauser, 1997). Just as individuals can have recognizable visual features, vocalizations can contain identifiable signature information. For example, individually specific identity information is often necessary for mothers to recognize their young so that females do not waste energy on raising unfamiliar offspring and to maximize

the chance that their offspring will survive through the period that they are incapable of nourishing themselves (Scherrer & Wilkinson, 1993; Phillips & Stirling, 2000). Signature vocalizations have also been recorded amongst groups of individuals. Sharing vocalizations between related kin may help to coordinate behaviors, such as foraging movements or caring for young, while also avoiding inbreeding between family members (Price, 1999; Christal & Whitehead, 2001). The social level at which these signatures appear could relate to the importance of a specific social bond. For animals that depend specifically on another individual for survival, individual signatures are often present while animals that benefit from being recognized as part of a group are more likely to benefit from sharing vocalizations characteristic of that group. On either level, these signatures must be unique enough not to be confused with another signature and distinguishable enough to be perceived after transmission through the medium. Ecological demands drive the evolution of a particular type of social structure and then this social structure will determine the individual relationships that occur. As vocal signatures demonstrate the dependence on particular interactions within a species, the amount that they correlate with different types of social structure reveals the significance of the social bonds within the social structure.

Vocal trends in individual-based societies

Individual signatures are present when inter-individual variation is greater than intra-individual variation for the same behavior (Falls, 1982). In general, animals that need to reunite with specific individuals in visually limited

environments will use calls that are more similar within an individual than between. The calls of highly colonial species have individual signatures which to help mothers to localize offspring within a large group. Loesche (1991) compared the amount of individual variation between cliff swallows (*Hirundo pyrrhonota*) and barn swallows (*H. rustica*) and found that the colonial species- cliff swallows- had more individually distinctive calls and argued that this is because barn swallows can use nest location to facilitate reunions. A similar experiment showed that northern fur seals (*Callorhinus ursinus*) have less within-individual call variation than northern elephant seals (*Mirounga angustirostris*) because fur seals remain close to their young offspring until weaning (Insley, 1992). Elephant seals do forage during the calving and breeding season, so females of this species need to localize their pups within a breeding colony and an individually unique and readily recognizable call would facilitate these reunions. Dealing with a similar but equally as crucial situation, many bat species use individually recognizable vocalizations as mothers must locate their offspring within a colony of many individuals or the offspring will die (evening bats, *Nycticeius humeralis*, Scherrer & Wilkinson, 1993; pallid bats, *Antrozous pallidus*, Brown, 1976; and Mexican free-tailed bats, *Tadarida brasiliensis mexicana*, Balcombe & McCracken, 1992; Balcombe, 1990).

Vocal trends in group-based societies

Sharing vocal signatures with other individuals in a group has been proposed in several types of situations. Because individuals of the same species

separated by physical distance do not interact, they do not share vocalizations and can develop different levels of vocal variation due to the geographic separation between groups. Within cetaceans, geographic variation within the vocalizations of female sperm whales separated by large distances of ocean has been reported (Weilgart & Whitehead, 1997). Wild chimpanzees are also thought to exhibit such variation between populations (Mitani *et al.*, 1999; Mitani *et al.*, 1992). Green (1975) found differences in call types used by three separate populations of Japanese macaques (*Macaca fuscata*). Geographic variation is found in the repertoires of bearded seals, *Erignathus barbatus* (Cleator *et al.*, 1989) and Weddell seals, *Leptonychotes weddelli* (Thomas & Stirling, 1983). Many bird species with large geographic ranges also show group-specific vocalizations (review in Mundinger, 1970). Group variation that is caused by geographic separation does not explain anything about interactions between groups and therefore sheds little light on social structure among interacting individuals.

In greater spear-nosed bats (*Phyllostomus hastatus*), social groups of unrelated females share screech calls as a type of membership badge which excludes non-group members from foraging trips to productive feeding areas (Wilkinson & Boughman, 1998). Australian magpies (*Gymnorhina tibicen*) have also been shown to share more songs with group mates and possibly use these group-specific vocalizations to defend a territory against neighbors identified by their non-shared vocalizations (Farabaugh *et al.*, 1988). Flocks of unrelated budgerigars (*Melopsittacus undulatus*) share group-specific social calls and new group members have been shown to modify their calls to match the group

(Bartlett & Slater, 1999). Stripe-backed wrens, *Campylorhynchus nuchalis*, also appear to be able to distinguish between the calls of their group and those of unfamiliar and unrelated groups (Price, 2003; Price, 1999). Playback studies have demonstrated a similar level of discrimination on the basis of group-specific vocalizations in Mexican jays, *Aphelocoma ultramarina* (Hopp *et al.*, 2001). These observations support the possibility that individuals of all of these species may derive some advantage from being identified to a group and social interactions are significant enough to develop a group distinctive vocal signature rather than an individual one.

Vocalizations in Cetacean Societies

Vocal patterns in bottlenose dolphins

The interactions between individual bottlenose dolphins have been well studied and are typical of a fission-fusion society regardless of population or geographic area (Connor *et al.*, 2000b). Bottlenose dolphins use individually specific vocalizations to reunite separated individuals when other cues are inadequate. These signature whistles have been repeatedly recorded prior to a reunion between isolated individuals both in captivity (Caldwell & Caldwell, 1965; Janik & Slater, 1998; Tyack, 1986; Janik *et al.*, 1994) and in the wild (Smolker *et al.*, 1993; Sayigh *et al.*, 1998). The distance that these whistles can travel demonstrates the ability for animals out of visual contact to communicate (Janik, 2000). However, within these societies there are two highly important interactions, both of which demonstrate increase vocal similarity as individuals

share signature whistles. First of all, calves depend on their mothers for survival and must maintain contact even though the demands of foraging require them to separate. In most bottlenose dolphin interactions, signature whistles are used to reunite, but mothers and calves have been shown to match signature whistles (Smolker *et al.*, 1993; Sayigh *et al.*, 1998). Similarly, male dolphins may experience increased mating success when partnered with another male to form an ‘alliance’ (Connor *et al.*, 2001) and male alliance partners also have matched signature whistles (Cook *et al.*, 2004; Watwood *et al.*, 2004). A high level of vocal similarity reflects the significance of these interaction between individuals.

Vocal patterns in sperm whales

Sperm whale vocalizations consist of various types of clicks, each with a different function. They produce regularly spaced echolocation clicks to scan for prey at depth and sped-up clicks known as ‘creaks’ when prey are encountered. Sounds known as ‘codas’ have been recorded from groups of whales socializing at the surface (Weilgart & Whitehead, 1993) . Coda repertoires appear to be group-specific as different individuals from the same group share a specific pattern (Rendell & Whitehead, 2004). Vocal similarities have been observed between groups of sperm whales and these units have been grouped into vocal clans based on the similarity of their vocalizations regardless of geographic range (Rendell & Whitehead, 2003). At high latitudes, males use a different pattern of clicks, termed ‘slow clicks’, which are rare at lower latitude (Madsen *et al.*, 2002). These slow clicks may allow dispersed individuals to interact as they

could have a greater range than coda vocalizations, although no studies have clearly documented this (Whitehead, 2003). Once again, vocal behavior and social interactions between individuals demonstrate a similar group-level trend. Group-based social structure among sperm whales is correlated with the presence of group-specific vocalizations for social females and a less group-specific vocalization for males that do not interact as a group.

Vocal patterns in transient killer whales

Some information already is available about the relationship between vocal behavior and social structure in killer whales. The difference in social structure between the two sympatric populations in British Columbian waters provides a potential correlation with contrasting levels of vocal behavior. While resident whales are highly vocal, transient whales vocalize much less and most frequently immediately after a kill (Barrett-Lennard *et al.*, 1996; Deecke, 2003).. Echolocation usage also differs between transient and resident animals and could reflect the fact that the seal prey of transient killer whales have sensitive hearing to these sounds (Barrett-Lennard *et al.*, 1996). Transients produce irregular, short, and slow clicks on rare occasions and, even when prey is absent, do not vocalize continuously, as clicks may reduce the possibility of opportunistically locating prey. Conversely, residents use regular, frequent echolocation clicks when foraging (Barrett-Lennard *et al.*, 1996). The higher level of acoustic sensitivity in marine mammal prey is thought to have caused transients to use passive listening instead of active echolocation to localize prey as seals avoided

the source when played the calls of transient whales (Deecke *et al.*, 2002). Deecke (2003) catalogued the repertoire of transient whales and found that transients have little variation in their vocal repertoire. Since these whales do not appear to have a social structure that is based on interactions between groups, there has been less selection for group-specific vocalizations. Individual interactions do seem to be important, so individuals may demonstrate much more individually specific vocalizations than within resident pods although this has yet to be investigated.

Vocal Patterns in resident killer whales

Many studies have documented the vocal similarities found within the stereotyped calls of resident killer whales at each level within their social structure. All killer whales produce three types of vocalizations, each related to a different context; echolocation clicks, whistles, and pulsed calls (Ford, 1987; Thomsen *et al.*, 2001). Echolocation clicks are short, broad band pulses used most frequently when foraging (Barrett-Lennard *et al.*, 1996) and variable whistles are narrow band tones used mostly during social interactions which appear to play a role in close-range interactions (Ford, 1989; Thomsen *et al.*, 2002). However pulsed calls, the third type of vocalization, are the most frequent vocalizations used by killer whales, and can be further divided into three types, all connected with social interactions. Pulsed calls can either be variable, aberrant or stereotyped. The first two are inconsistent in frequency structure and rarely used by the whales (Thomsen *et al.*, 2002). The following discussion of vocal

repertoire relates only to the study of the stereotyped pulsed calls, which can be identified through their consistent and stereotyped structure. They have been classified through a characteristic set of components, or repeatable changes in the pulse repetition rate over time, and can be easily discriminated from each other by visual inspection of spectrograms (Ford, 1987; Ford, 1989). Ford (1987) classified these stereotyped calls based on their 'call type', or the consistently recognizable sequence of components, and catalogued them using a letter to describe their geographic population ('N' for northern community, 'S' for southern community and 'T' for transient individuals) followed by a number to describe the type. While all call types contain the lower frequency component that ranges from 80 to 2400 Hz (hereafter called the 'low-frequency component' or 'LFC'), some types have a highly identifiable 'double' structure; these calls consist of a LFC combined simultaneously with a second component (hereafter called the 'high-frequency component' or 'HFC') with a fundamental frequency from 2 to 12 kHz (Hoelzel & Osborne, 1986; Miller & Bain, 2000).

Groups of individuals that associate have been shown to use the same set of call types. Based on this observation, the sixteen pods of resident whales have been divided into four vocal clans that do not use any of the same call types. Within vocal clans, pods share some call types but not all of them, and exhibit slight but consistent variations in the stereotyped structure (Ford, 1991; Ford, 1987). These variations have also been demonstrated in the smallest group within the resident social structure, the matrilineal unit, and are distinct enough to distinguish between matrilines in the same pod (Deecke, 2000; Miller & Bain,

2000). A similar sharing of call types has been reported for both Alaskan and Norwegian populations (Yurk *et al.*, 2002; Strager, 1995). Ford (1989) proposed that group-specific vocalizations maintain contact within a pod and delineate membership status through the use of shared call types. Slight errors in call copying between generations may have progressively caused structural differences to appear between matrilineal units and over time to accumulate into significant differences between pods (Deecke, 2000; Miller & Bain, 2000). Later genetic studies have supported the correlation between shared repertoires and genetic relatedness (Yurk *et al.*, 2002). However, acoustic differences do not match the observed social associations of individuals, as pods from different clans have been observed traveling together (Ford, 1991). Individuals of the same matrilineal unit travel together most of the time and consistently interact and hear the vocalizations of their family members. If temporary associations, such as those between matrilineal units, have produced consistently recognizable vocalizations, the possibility arises that permanent associations also have an impact on vocal structure. Some evidence for individually specific vocalizations has been documented (Hoelzel & Osborne, 1986), but the relative relationship between such signatures present at the level of the individual and those at the level of the matrilineal group has not been explored. The strength of social bonds between groups of individual killer whales suggests the possibility that these individuals only need to communicate group-specific information, therefore making the identification of individuals within that group unnecessary. This study explores whether the social structure of resident killer whales has selected

for the communication of only group identity and whether vocal interactions therefore lack the presence of any individual information.

CHAPTER 2:

Neural networks for pattern recognition to discriminate between calls of resident killer whales with different levels of social association

Data collection

Recording individuals in the field

During August and September 1998 and 1999, recordings were made of northern resident killer whales inhabiting Johnstone Strait, British Columbia. Following the methodology developed by Miller & Tyack (1998) and used by Miller *et al.* (2004), groups of whales were tracked from an 11-meter research vessel using the focal follow technique described by Altmann (1974). Recordings were made with a towed array of 16 hydrophones connected to multi-channel TASCAM recorders, which digitally recorded sounds at a sampling rate of 48 kHz. Additionally a single channel from the array was recorded at a 96 kHz sampling rate on a Pioneer D-9601 DAT recorder, on which voice notes described the position of the focal animal and the location of other animals in the area (Miller & Tyack, 1998).

Individuals were identified through recognizable dorsal fin shapes and saddle patches that have been compiled into a photo-identification catalog by Ford *et al.* (2000). While recording on the hydrophone array, the vessel was maneuvered to attempt to keep the focal animal directly broadside of the array at a distance of approximately 100 meters, a position that minimizes the effect of depth on angle-of-arrival (Miller & Tyack, 1998). The position of the focal animal relative to the vessel was recorded at each surfacing with a laser range finder and

digital compass. Focal position between surfacings was estimated by interpolating surface positions and including a correction factor to account for depth (Miller & Tyack, 1998).

Ascribing calls to an individual - The Beamforming Technique

Raw field recordings were transferred to a computer with Cool Edit Pro and then processed with custom MATLAB software to create a FFT spectrogram (with an effective filter bandwidth of 94 Hz and a dynamic range of 50 dB) synchronized with a directogram from the calculated angle-of-arrival of recorded calls. The angle-of-arrival was calculated using the broad-band frequency domain beamforming algorithm developed by Miller & Tyack (1998). This algorithm found the angle-of-arrival of vocalizations through measuring the differences in energy at different angles of arrival after correcting for expected frequency phase shifts. This steered response power was plotted over time to generate the directogram, and identifies the angle with the most power at each frequency.

Coordinating the output of the beamforming algorithm with observed animal locations relative to the array, specific calls could be ascribed to individual whales. To assure that calls beamformed to a specific direction were produced by the animal observed at that location, only calls produced when the focal individual was separated by 20° from all other animals were selected from the continuous sequence. Spectrograms of each isolated call were made in MATLAB and inspected visually for classification within the catalogue of stereotyped resident killer whale call-types published by Ford (1987). All vocalizations that

were not clearly classifiable as stereotyped pulsed calls were labeled as ‘variable calls’ and were not analyzed further.

Time-frequency contours

Before similarity can be measured, a representation of the acoustic signal must be created that accurately represents each recorded vocalizations. For this study, frequency contours were used to track the pulse-repetition rate of each vocalization, as frequency contours have been determined to be robust in representing the calls of killer whales (Deecke et al., 1999; Deecke, 2000) and the whistles of bottlenose dolphins (Caldwell & Caldwell, 1965; Buck & Tyack, 1993; Ralston & Herman, 1995; McCowan, 1995). This method allows a signal to be analyzed as a whole unit, a representation that is more biologically relevant than the measurement of discrete parameters, as in Ford (1987), Hoelzel & Osborne (1986), and Miller & Bain (2000). Additionally, the comparison of complete frequency contours eliminates any subjectivity involved with selecting parameters to include.

In general, a time-frequency contour models the frequency modulations over the duration of a vocalization. While several algorithms have been used to extract contours from the overall spectral content of cetacean vocalizations (Buck & Tyack, 1993; Deecke *et al.*, 1999), this study uses the pitch detection algorithm (PDA) developed by Wang & Seneff (2000) as it had two main advantages over other methods. First, it allows the frequency to be continuously tracked over the course of the signal, unlike other algorithms which measure the energy of the

pulse-repetition rate at discrete intervals during the signal duration. Secondly, it searches across a specified range of frequencies to locate and sample the spectral peaks that correspond to the harmonics of the fundamental frequency, allowing contours to be extracted from signals with a masked fundamental (Wang & Seneff, 2000). Unfortunately, this algorithm must identify a frequency value for every point of a signal, so that portions without frequency content will also have an extracted value. Also, the extracted frequency may be doubled or halved. To reduce the influence of these disadvantages, the two additional steps detailed below were performed on contours after extraction.

This algorithm combines two functions to track the fundamental frequency across the signal. The first function correlates a template of the spectral peaks across several frequencies by logarithmically sampling multiple harmonic peaks to locate the fundamental frequency. The second function then calculates changes in the fundamental across those adjacent points. By combining these two functions, the algorithm continuously tracks the fundamental frequency across the whole signal duration, instead of only at discrete frequency points (Wang & Seneff, 2000). This algorithm has been tested against other pitch detection algorithms used for detecting human speech and has demonstrated significantly better results (Wang & Seneff, 2000).

Contour extraction

Frequency contours of both low and high frequency components were extracted separately using the above-mentioned pitch detection algorithm developed by the Massachusetts Institute of Technology Spoken Language Systems Department. Extracted contours were double-checked visually to mitigate the two main drawbacks of this algorithm; requiring a value regardless of spectral content and the possibility of pitch doubling or halving. This correction was performed with custom MATLAB software which overlaid each contour onto a FFT spectrogram of the signal. Beginning and ending points were selected to eliminate all points that corresponded to regions without pitch content. Finally, any contour points that did not line up with the spectrogram were adjusted to remove the occasional pitch doubling. The final extracted LFC contour was matched with the second frequency harmonic of the lower frequency portion of the signal and the HFC contour was matched with the fundamental frequency of the higher frequency portion shown on the spectrogram.

From this point on, the LFC and HFC contours were dealt with completely separately. After correction, each contour was rated on a scale of 0-100 to classify the percentage that was clearly visible. Contours that were rated one hundred percent complete were interpolated to a standard length of 100 frequency points. Call duration was added to the frequency vector in order to differentiate between calls of similar shape but different length, making the total vector length 101 frequency points (Deecke *et al.*, 1999). These standardized calls were then ready for comparison using a neural network pattern recognition program to

analyze the level of similarity between calls at four different levels of the killer whale social structure.

Measurement of call similarity

Neural network pattern recognition

Since a multi-layer feed-forward network with supervised back-propagation learning has already been demonstrated to be effective in the inter-group discrimination of killer whale vocalizations (Deecke et al., 1999), it was the network structure used in this analysis. For this study, 'neural network' will hereafter be used to describe only this particular architecture, even though it is not the only form of neural network analysis that is possible. However, it is important to remember that there are many types of neural network architectures effective in both pattern recognition and other forms of clustering analysis.

Network structure

The basic goal of the network is to classify a signal regardless how fine-scale the difference in the characteristics. To do so, the network will compare an unclassified input signal with previously defined signals and determine the pattern to which the input is most similar. Therefore, a reference pattern must be developed through a process called 'training the network' before a signal can be tested and identified. One of the most significant advantages of neural network classification is its ability to utilize several examples in generating a reference pattern, allowing the inclusion of natural variation and creating a more robust representation based on all available external information (Hinton, 1992). The

network does this through processing inputs via hierarchical layers, where information from several inputs is collected, summarized and then passed on to another layer for combination with the summarized information from other sources. Such a layered system theoretically models the way in which sensory information is collectively processed within the nervous system (Hinton, 1992).

To train the network, each reference example is multiplied by a randomly generated value known as the ‘synaptic weight’ and added to the reference pattern. The amount of error generated by the addition of a new example, δ_{pj} , is calculated using the following general delta rule, where $f'_j (net_{pj})$ is the derivative of the function relating the input set to the output set, δ_{pk} is the previous amount of error and w_{kj} is the value of the previous synaptic weights.

$$\delta_{pj} = f'_j (net_{pj}) \sum_k \delta_{pk} w_{kj} \quad (\text{Rumelhart } et al., 1986)$$

The synaptic weights are then modified until the error term is below a previously determined threshold known as the activation function. Different activation thresholds allow for more complexity in resolving specific situations; for example, distinguishing between more closely related patterns (Rumelhart *et al.*, 1986).

Once a reference pattern has been developed, the ‘testing’ stage of the network comparison begins. New signals are presented to the first layer of the network, known as the input layer. The characteristics that define the reference pattern are contained in the internal network layers, layers which are known as hidden neurons (Rumelhart *et al.*, 1986). If an unknown signal corresponds to the network’s internal representation, the new signal will be identified as that pattern.

The classification 'decision' will be returned in the final layer of the network, which is also known as the output layer.

Neural network justification

While several different methods of measuring acoustic similarity are available, the use of neural network performance to assess similarity presents a more objective and repeatable alternative to these established techniques. In one method, several human observers are asked to match more similar sounds based on visual inspection of their contour and aural discrimination of the sound. The scores of multiple observers are combined to create a similarity score for a series of calls (Deecke *et al.*, 1999; Yurk *et al.*, 2002; Watwood *et al.*, 2004). Although this method allows a high level of flexibility in identifying similarity in slightly different calls, it is highly subjective and will result in different similarity scores from different observers. Computer-based clustering algorithms have the advantage of producing repeatable results, but run into difficulty when faced with the slight and inconsistent variations common in biological settings (Janik, 1999). Studies that test the results of computer classification against those from human observers found that while computer methods are more consistent, human observers can generally perform better (Janik, 1999; Watwood *et al.*, 2004). Through the use of a training set to develop a reference pattern, the network accounts for natural vocal variations. Comparison between the similarity measured by a neural network algorithm and that of human observers produced similar results (Deecke *et al.*, 1999).

Although network comparisons have the advantages listed above, they are restricted by several disadvantages. First of all, network performance depends on the examples supplied in the training set with both the quantity of training examples as well as the proportion of each patterns within those examples influencing network discrimination. Networks will discriminate between test patterns more accurately when more examples are included in the training set (Rumelhart *et al.*, 1986). Additionally, equal numbers of each example must be included in the training set to avoid a weighted network discrimination (Rumelhart *et al.*, 1986). The following portions of this chapter further describe the methodology used to reduce these drawbacks. Finally, network comparisons are limited by their ability to identify the fine-scale features on which classification is based. While the identification of distinguishing features is necessary at some point, the general discrimination of individual calls will be sufficient for this initial study.

Regardless of these shortcomings, this method was chosen based on its ability to reduce the objectivity inherent in human classification techniques, and also include complete vocalizations. Previous studies of killer whale vocal similarity have shown it to be robust in comparing frequency contours, and therefore demonstrate adequate performance in this study.

Comparison bootstrapping

Following the protocol developed by Deecke *et al.* (1999), pairwise comparisons of calls from two individuals were performed with custom

MATLAB software involving the Neural Network Toolbox. In order to confirm the accuracy of the network results, this program used a bootstrapping procedure, where one call was removed from the subset and the network trained on the remaining calls. Since these were comparisons between the calls of only two individuals, all novel patterns could be classified to one of two possible contours. Each call in turn was used as the novel contour and was classified as one of the two possible patterns generated from the other calls.

During training a binary target output was used to describe the two possible patterns; [1 0] represented calls from one individual while [0 1] represented calls from the other. One hundred fifty training iterations were then used to associate each pattern with its target output. Finally the excluded call was presented to the network and classified as belonging to one of the two patterns. Deecke *et al.* (1999) found that accurate performance of the network required 20 hidden neurons and 5000 training iterations in order to perform pairwise comparisons between matrilineal units. After completing 5000 classification iterations, the program calculated the average discrimination error (ADE) by comparing the absolute mean difference between observed and expected outputs.

Selection criteria for network input

Network performance can be influenced by several aspects of the system, regardless of the data for analysis. Each network must be generated from equal numbers of input patterns. When more examples of each pattern are used within the training set more natural variation is included in the network, allowing the

network to better classify similar patterns containing biological noise. The earlier study comparing the stereotyped calls of killer whales used the smallest number of calls available for one matrilineal unit (24 calls) as the input size for all comparisons. Since many matrilineal units had more calls ascribed to them, calls were selected from across the maximum number of recording sessions and behavioral states. Any calls in excess of the 24 specifically chosen calls were then excluded from further analysis (Deecke *et al.*, 1999). Therefore, in the present study trials were needed to find the number of hidden neurons, and the minimum number of calls necessary while maximizing the number of comparisons.

Input size trials

While Deecke *et al.* (1999) described the optimum number of hidden neurons and training iterations, neural network analysis still requires roughly equal amounts of each input pattern to generate two cumulative representations for comparison. Due to the difficulty in obtaining field recordings, the large number of individuals involved, and the inability to record specific vocalizations, sample sizes were very uneven across individuals as well as call types. Therefore, selecting a large input size would reduce the number and type of comparisons possible as fewer animals had that many calls ascribed to them. The social affiliation of individuals was also considered when selecting the input sample size as comparisons at different levels of social affiliation were needed. The self and within-MU comparison types required a smaller sample size, as sufficient

amounts of calls were not available for more than one member of several matrilineal units. The comparison of different calls from the same individual limited sample size the most to a maximum input size of 20 calls. Additionally to perform a within-MU comparison, twenty calls had to be available for two individuals from the same matrilineal unit and this was possible for only one matrilineal unit. The input size was reduced until a broader representation of self and within-MU comparisons was possible, finally settling on an input size of 15 calls from each individual (30 calls/comparison). This sample size resulted in a total of 37 inter-individual comparisons of a shared call type (Figure 2.1) and allowed 18 between-MU, 9 within-MU and 8 self comparisons to be performed.

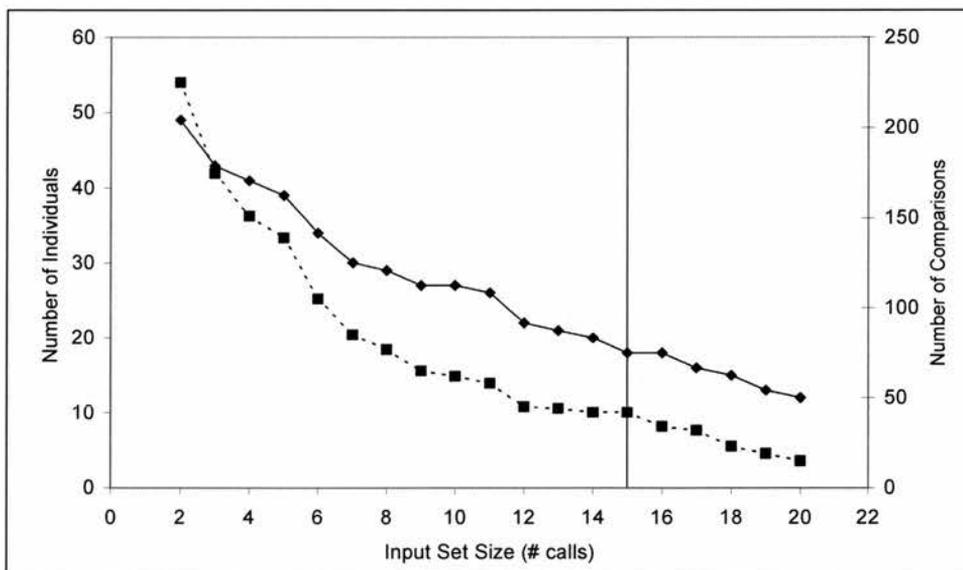


Figure 2.1: Diagram of relationship between input size and number of possible comparisons. Solid line shows number of individuals included for a particular input size, dashed line shows number of comparisons possible using that number of calls. Vertical line shows chosen input size for selected comparisons.

Randomization Trials

The process of selecting individual calls for inclusion in the analysis excluded many individuals for whom data were available from the study as well

as reducing the amount of variation that would exist within natural interactions. In order to incorporate all variation into the analysis, increase the number of comparisons possible (through increasing the number of individuals with more calls than the input size required) and select the input sets without any observer bias, a technique of randomly selecting input sets was developed using custom MATLAB software.

In order to test the randomization software, the network algorithm was run with input sets of 4,6,8,10 and 15 calls selected at random from the three individuals with the most calls available (A32, A46 and A37). These trials were then evaluated on the similarity of results between input set sizes. The eight call set was found to generate similar results to the largest possible input size (15 calls) so it was selected as the minimum input size for further network comparisons, and using this input set size made sixty-eight inter-individual comparisons possible (see Figure 2.2 for trials with one comparison).

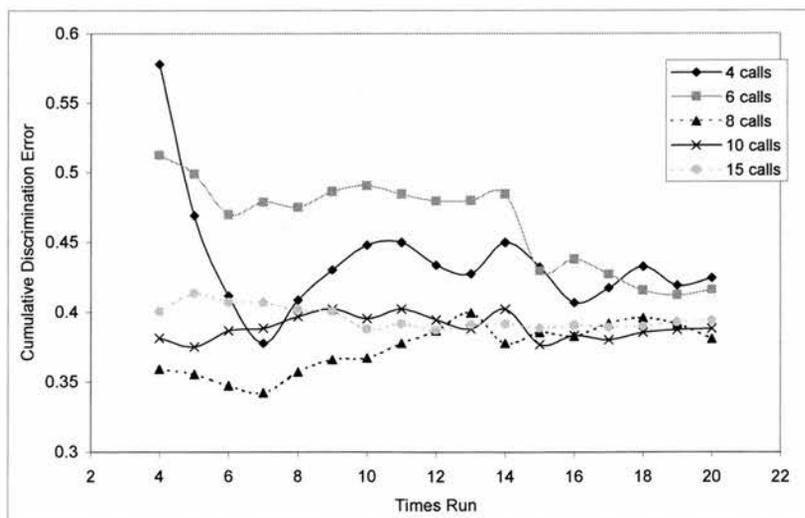


Figure 2.2: Results of input size trials with comparison of N4 calls from individuals A46 and A32. Plot shows network discrimination error averaged over the number of times the network was run (from 4 to 20) with various amounts of calls for each individual. Grey dashed line with circular points shows network results for maximum input size, black dashed line with triangular points shows network results for target input size.

The number of randomizations necessary to include all available calls was evaluated through plotting the cumulative average of the network output for 4 to 50 randomizations of 8 calls for comparisons involving individuals A32, A37 and A46. An input size of 4 calls was also tested over this increased number of randomizations to test if including more calls even with a smaller sample size influenced discrimination error (Figure 2.3). The cumulative average of network results reached an asymptote after approximately 30 randomizations for all comparisons performed between the three individuals, and eight calls was determined to produce optimal network performance.

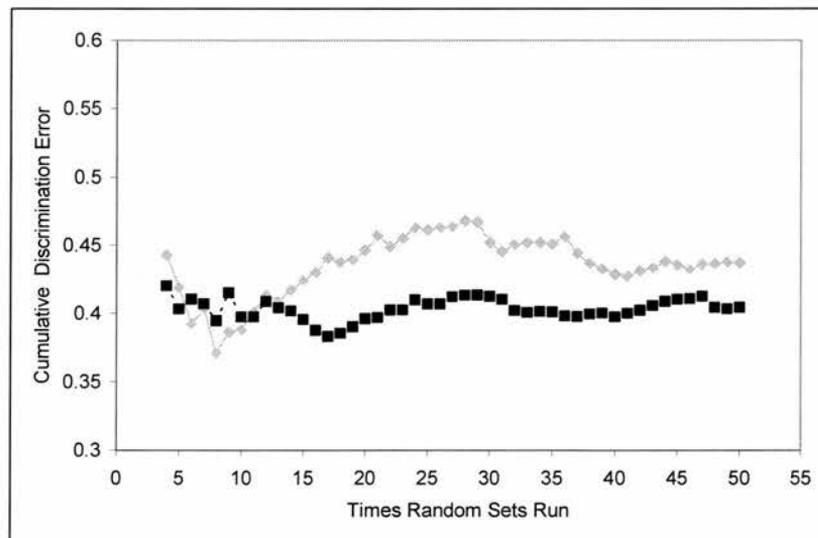


Figure 2.3: Results of randomization trials for comparison of N4 calls from individuals A46 and A32. Trials performed through running network with randomly selected sets of four calls (grey line with diamond points) and eight calls (black line with square points) and averaging the resulting discrimination error over the number of runs.

Conclusion

Due to the high level of sociality of these cetaceans, previous studies of acoustic similarity have never been able to ascribe calls to the individual that

produced them in a group. Although recordings can be isolated to an individual in captivity, the calling behavior of such animals is likely not indicative of natural communication and behavioral interactions in the wild. Previously, there have been several successful comparisons of killer whale group vocalizations. For the first of these studies, Ford (1987) used discrete statistical comparisons of call components to quantify the differences between call types. Miller & Bain (2000) used a discriminant function analysis to quantify individual call features, while Deecke *et al.* (1999) used both a neural network algorithm and human observers to evaluate call similarity. Because they have been shown to be robust with killer whale vocalizations, and emphasize the biological context of vocalizations, neural networks are useful for an initial and general analysis of overall similarity in killer whale vocalizations. The goal of this study is to apply the comparison method of Deecke *et al.* (1999) to the calls recorded with the beamforming technique of Miller & Tyack (1998) in order to generate a detailed comparison of calls from individual whales, a step which has previously not been possible.

To assess the significance of vocal behavior within killer whale social structure, this study included comparisons at different levels of affiliation within the resident social structure. Based on available data, four levels of social affiliation were defined and comparisons were performed between individuals at these levels (hereafter known as comparison types). The types were comparisons between (1) one individual (self comparison), (2) two individuals within the same matrilineal unit (within MU comparison), (3) two individuals from different matrilines (between MU comparison). Additionally, three repertoire levels were

used: (1) calls of the same type (within call type comparison), (2) different call types from a repertoire where all call types were shared (within shared repertoire comparisons) and (3) different call types from repertoires with no call types in common (between unshared repertoires).

To confirm that the level of similarity measured between calls of individuals was not an effect of call selection, three experiments will be used to find the method of call selection which best measures vocal similarity. The first explored any differences between selecting specific calls according to context and randomly selecting them by computer, using low-frequency component contours of the same call type as a test set. The second experiment is designed to investigate the possibility that conspecific vocalizations influence the similarity measured between calls from the same individual. During this experiment, the similarity of calls chosen irrespective of calling sequence will be compared with the level of similarity between calls from entirely separate calling sequences, also only looking at the low-frequency component within one call-type. These two experiments will indicate any variations that are effects of context and call selection but are not characteristic of individual signatures. The best form of call selection, as identified from the results of the two previous experiments, will then be applied to comparisons of the high-frequency component as well as comparisons between different call-types in the third and final experiment. Through identifying the extent of vocal signatures within the social structure of killer whales, this study attempts to better understand the calling behavior of individual killer whales and its function within their social relationships.

CHAPTER 3:
*The investigation of optimal network selection to
assess the relationship between vocal signatures and
resident killer whale social structure*

INTRODUCTION

Contextual effects have repeatedly been shown to influence the calling behavior of killer whale groups. Therefore, before the presence of individual signature information within the stereotyped calls of northern resident killer whales can be confirmed, other possible influences on intra-individual variation must be tested to determine the extent to which they affect the analysis. This experiment tested whether the method of selecting calls to be compared has any effect on the level of similarity measured in order to develop an optimal network procedure. Neural network optimal procedure was measured by running the same comparisons with calls selected in two different ways. Firstly, calls were selected individually across as many recording sessions as possible. Secondly, a smaller set of calls was selected randomly by computer and then compared by the network. After a series of these sets were compared, the resulting level of similarity was averaged. Grouping the results from either procedure on the basis of social relationship between the individuals involved enabled the comparison of intra-group, inter-individual and intra-individual similarity levels.

Bain (1986) found an association between an increase in how often killer whales call and raised arousal, suggesting that more vocalizations would signal the location of a prey patch to conspecifics. Ford (1989) measured distinct changes in call length that he attributed to increased intensity of social

interactions. Although no correlation between specific call-types and behaviors has been shown (Ford, 1989; Morton *et al.*, 1986), vocalizations may be more frequent during active behaviors such as foraging and milling than during resting behaviors (Hoelzel & Osborne, 1986). To be certain that calls were produced by a specific individual that focal animal had to be separated from the group (see chapter 2), making the inclusion of specific behavioral contexts difficult in this study. Generally focal animals were traveling or foraging apart from the group, although other behaviors were occasionally observed and noted (see Appendix 2). Features that indicate individual identity should be consistent across contexts and different days as intra-individual variation should be less than inter-individual variation (Falls, 1982). Therefore even though specific behavioral data were not available, including calls from the largest number of possible calling situations is necessary to strengthen the conclusion that features identified within the analysis are consistent because they are from the same individual and not just from the same context.

Comparing calls with a neural network allows the generation of a reference pattern representative of several examples. Although this is very useful with biological data to reduce the influence of individual variation, this method of comparison adds an area of uncertainty. Inappropriate selection of example patterns can influence network performance and therefore place several constraints on the procedure of running the network. The two most limiting requirements of the network are 1) a large number of inputs and 2) equal amounts of each input (Rumelhart *et al.*, 1986). The difficulties of field recording,

especially with cetaceans, make these limitations very restrictive. In the present study, very few vocalizations were recorded for some individuals, making those individuals impossible to include within the analysis (see Chapter 2). A smaller input size would allow more individuals to be studied, but also would reduce network performance. Additionally, as the size of inputs from both individuals being compared must be the same, the smallest number of calls from one individual dictates the number of calls that must be used in every comparison. Any further calls from other individuals would be ignored in the analysis. The need for selection also creates a subjective element in the analysis as some selection criteria determine which calls to compare within the network.

Deecke *et al.* (1999) selected calls from as many different recording sessions and behavioral contexts as possible in order to control for any effects of emotional state. Their comparisons were performed with an input size of twenty-four calls per matrilineal group (forty-eight calls total per pairwise comparison). However the selection of calls from different sessions greatly reduced the proportion of the data used and also incorporated a subjective element into the analysis. For both of these reasons, the inclusion of all available data would be the best way to reduce the influence of situational variation and increase the chances that the observed differences are indicative of individual variation.

METHODS

Recordings were made of northern resident whales inhabiting the Johnstone Strait region of British Columbia during August and September 1998

and 1999. Using the focal follow technique described by Altmann (1974), groups of whales were followed in an 11-meter research vessel that towed a hydrophone array. Calls were recorded using custom beamforming software to determine the angle of call arrival on a hydrophone array and then ascribed to individuals by matching these angles with behavioral observations on the position of whales relative to the recording vessel (Miller & Tyack, 1998). Contours were extracted from the spectrograms using the pitch tracking algorithm developed by Wang & Seneff (2000), corrected manually and input into the neural network developed by Deecke et al (1999). Calls were identified to call type using the naming system of Ford (1987). All comparisons were between calls of the same type. See Chapter 2 for further details.

Comparisons were grouped into the following three types according to the social affiliation of the individuals; between matrilineal unit (MU), within matrilineal unit (MU) and self. The social associations published in Ford *et al.* (2000) were used for identification of individuals in the field, so they will be used to classify the affiliation levels for comparisons made in this study. Animals for which calls were available are members of two clans, A and R clan, but it is important to remember that there are a total of four clans, with a third clan belonging to the northern community, and a fourth that makes up the southern resident community (Figure 3.1).

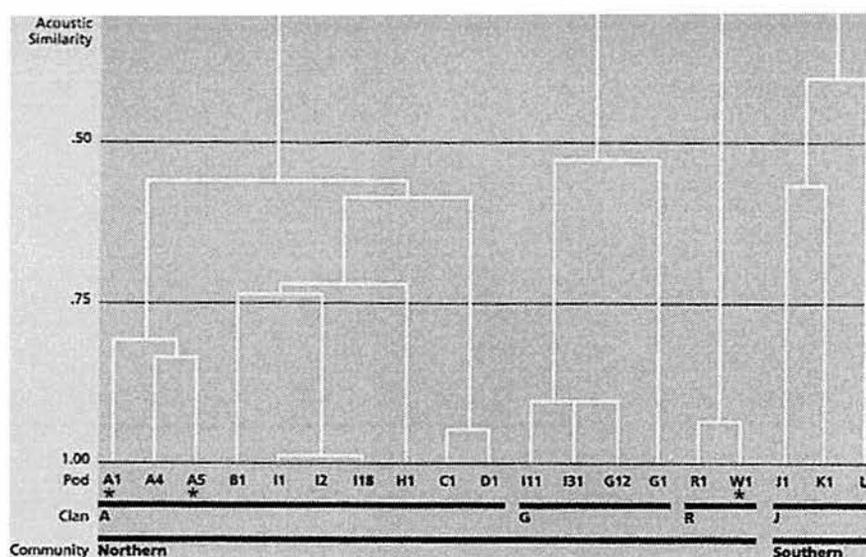


Figure 3.1: Social affiliation of killer whale pods from Ford *et al.* (2000). Pods are listed across the x axis and divided by clan and community, while increasing acoustic similarity is shown on the y axis. Pods delineated with a star (*) are those from which calls were recorded.

A clan is the largest of the four clans, with a total of 10 pods while R clan consists of 3 pods. Clans were recorded from 19 members of three of these pods, A1, A5 and W1. Pod A1 consists of three matrilineal units (A36, A12 and A30), while pod A5 has four matrilineal units (A8, A9, A23 and A25) and W1 is formed from a single matrilineal unit (W3). The relationship of individuals within the matrilineal units for which calls were recorded will be described below in the relation to the available dataset (see page 3-8).

Fifteen calls per individual were found to be the ideal input size for the hand-selected procedure (see Chapter 2). Therefore for each individual in a pair to be compared, fifteen calls of the same call type were selected from the maximum number of recording sessions possible and over the widest temporal separation within recording sessions. For self comparisons, two sets of 15 different calls were selected from those available for a single individual, requiring a minimum of 30 calls in order to perform this comparison. The frequency

contour vectors of all 15 calls were combined into a single matrix, and the target outputs, which indicate the individual to which the contour corresponds, were put into a second matrix. Both the contour matrix and target matrix were then presented to the network program.

Alternatively, to compare call structure using calls of the appropriate call type selected at random from all of those recorded from an individual, a MATLAB program was created to select eight calls of the same call type from each of two individuals. A series of trials had previously shown that inputs sets of eight calls randomly selected thirty different times from the set of all calls of the appropriate call type produced consistent network performance with this program (see Chapter 2). To generate random input sets, the frequency contours of the sixteen randomly selected calls were combined into one matrix, the corresponding target outputs compiled into a second matrix and the network run on this set of calls. For self comparisons, two sets of eight different calls were randomly selected from the total recorded for a particular individual. Thus, at least 16 calls had to be recorded from an individual for the self comparison to be made. When completed the result was stored, and a new set of eight calls from the same call type and same individuals was selected at random. This random selection procedure was performed a total of thirty times and the results of the network were averaged over the entire thirty runs.

The output for these networks is the average discrimination error (ADE) of the two calls and describes their level of similarity. After running every call through the network, the program averages the mean of the absolute difference

between the observed network output from the expected output, a modified version of calculating the proportion of calls correctly classified to individual against the number incorrectly classified (Deecke *et al*, 1999). Because the network is classifying each call to one of the two possible individuals, calls that are more similar will be more difficult to classify. A higher discrimination error means that more calls were incorrectly classified and since more similar calls are harder to classify, a higher ADE is produced by more similar calls.

To establish the discrimination error level that relates to discrimination by chance, additional comparisons were made between two sets of identical calls taken from the same recording session and individual. For these control sets, eight calls of the same call type were selected from the same individual during a single session and then combined with a second copy of the same eight calls to form an input set of two identical patterns. The results of comparing thirty randomly selected sets with the network were averaged as with further comparisons (see Chapter 2). As all factors analyzed in further analysis were limited for these comparisons, these call sets should establish the error level produced by chance discrimination of the network.

For each selection method, significance between comparison types and call types was measured with a three-way analysis of variance test. Significance levels of this test were also corrected with the Bonferroni adjustment to avoid errors generated from multiple testing (Sokal & Rohlf, 1981). However, this did not alter any of the comparisons made. Normality was confirmed using the Shapiro-Wilks test for normality and homogeneity of variance was determined

with Levene's test. Although two results of this test were significant, they were for factors with very small sample sizes. As all other results were strongly not significant, these results were attributed to the low power of these data. Therefore, the influence of selection method on similarity level was then compared parametrically (Zar, 1996). Finally, a Kruskal-Wallis test was performed to compare chance network performance with the discrimination error measured for randomly selected comparisons to confirm that measured discrimination differed significantly from chance. All statistical tests were performed with SYSTAT software (SPSS Inc).

RESULTS

Recording sessions from 27 different days were used, involving 72 separate focal follows and a total of 27 recording hours. These recording sessions yielded a total of 1508 vocalizations from 19 different individuals of A and W pods. Eight hundred calls could be ascribed to a specific individual and 724 of these calls were classified as one of the seven most common call types. Six hundred and thirty were A pod call types (N1, N2, N4, N5, and N9), while the remaining ninety-four calls were W pod call types (N32 and N33). The number of calls within each call-type dictated which call types were included in the following comparisons (Table 3.1). For hand-selected comparisons, this limitation resulted in input sets to be selected from a total of 491 available calls of five call-types (N2, N4, N5, N9 and N33) from the following seven individuals: A32, A46, A37, A6, A38, A8 and W2. The random-selection method allowed

more calls to be included in the overall sample set, as it required a minimum of eight calls per individual per call type. This sample set had a total of 626 calls of seven call-types (N1, N2, N4, N5, N9, N32 and N33) from twelve different individuals.

Pod / Matrilineal Unit	Individual	Sex	N1	N2	N4	N5	N9	N32	N33	Total
A1 / A36	A32	M	8	23	19	16	28	0	0	94
	A46	M	14	16	41	37	52	0	0	160
	A37	M	12	2	27	18	11	0	0	70
A1 / A12	A12	F	1	11	3	2	0	0	0	17
	A33	M	0	0	0	4	0	0	0	4
	A34	F	0	6	5	5	2	0	0	18
	A55	M	2	6	13	6	11	0	0	38
	A62	F	0	0	8	0	5	0	0	13
	A67	F	0	0	0	2	1	0	0	3
A1 / A30	A6	M	1	0	17	0	7	0	0	25
	A38	M	0	2	26	0	18	0	0	46
A5 / A8	A8	F	0	44	40	4	35	0	0	123
	A28	F	0	1	1	0	0	0	0	2
A5 / A23	A23	F	0	1	0	0	3	0	0	4
	A27	M	0	1	0	0	0	0	0	1
	A60	M	0	6	1	0	5	0	0	12
W1 / W3	W2	M	0	0	0	0	0	20	34	54
	W3	F	0	0	0	0	0	11	14	25
	W5	M	0	0	0	0	0	5	10	15
Total			38	119	200	94	178	36	58	724

Table 3.1: Sample size of calls recorded from northern resident killer whales separated by individual (rows) and call type (columns). Calls were recorded for 19 individuals (11 male, 8 female) of two pods and six matrilines. Horizontal lines divide individuals into matrilineal units, and bold horizontal lines group individuals by pod (see Figure 3.1 for relationships between pods).

Calls were recorded from individuals of six different matrilines within pods A1, A5 and W1 (see page 3-5 for the relationship between matrilines). All genealogies described here were taken from Ford *et al* (2000). All three matrilineal units of pod A1 had calls included in the dataset. For matrilineal unit A36, calls were available from all three members A32, A46 and A37. Notably, these individuals were all male, as their mother, A36, was deceased and therefore there were no further generations. Matrilineal unit A12 consisted of seven individuals, and the calls of six were recorded. The eldest female, A12, was still present along with her two male (A31 and A33) and one female (A34) offspring.

Additionally, A34 had three offspring, one male (A55) and two females (A62 and A67). The third matrilineal unit of A1 pod was A30 and had the following six members at the time of this study; the surviving female A30 and her five offspring, three of which were male (A6, A38 and A39) and two of which were female (A50 and A54).

From pod A5, two matrilineal units (A8 and A23) were represented in the dataset while the other two (A9 and A25) were not. There were four individuals in matrilineal unit A8; female A8, her two female offspring, A28 and A42, and one juvenile offspring of A42. The membership of matrilineal unit A23 consists of female A23, and her two male (A27 and A60) and one female (A43) offspring. Matriline A9 consisted of two males, A5 and A26, and matriline A25 included a female, A25 and her female offspring (A51). Pod W1 consisted of only a single matrilineal unit with only three individuals. This group consisted of a female, W3, and her two male offspring, W2 and W5.

Hand Selection Method – A pod

Only individuals representing two matrilines (A8, A32, and A46) had enough calls to measure the acoustic similarity of these individuals (Table 3.2). Averaging comparisons by comparison type ranked mean ADE from least to greatest: between MU, within MU and self comparisons (Table 3.7).

A8	<i>0.5717</i>							
A32	0.2128	----						
A37	----	----	----					
A46	0.2528	0.3166	----	----				
A12	----	----	----	----	----			
A55	----	----	----	----	----	----		
A6	----	----	----	----	----	----	----	
A38	----	----	----	----	----	----	----	----
N2	A8	A32	A37	A46	A12	A55	A6	A38

Table 3.2: Average discrimination error (ADE) for each pairwise comparison of N2 calls selected by hand from all available N2 calls from A pod members. Individuals being compared are listed in bold font in the left most column and bottom most row. Values generated from 15 vocalizations per individual (30 calls per comparison). Normal text/clear box represents the results of between-MU comparisons, bold font/light gray box indicates within-MU comparison results and italic font/dark gray box represents comparisons between calls from the same individual. Dashed lines indicate sample size was too small to perform a comparison.

The N4 call type had the largest number of calls, making a total of seventeen comparisons possible for the individuals of three matriline: A6, A8, A32, A37, A38 and A46 (Table 3.3). The three comparison types can be ranked in the following order from least to greatest mean ADE: between MU, within MU, and self comparisons (Table 3.7).

A8	----							
A32	0.1967	----						
A37	0.2707	0.3693	<i>0.5533</i>					
A46	0.1762	0.4401	0.3645	<i>0.5559</i>				
A12	----	----	----	----	----			
A55	----	----	----	----	----	----		
A6	0.0918	0.2363	0.2657	0.2809	----	----	----	
A38	0.0965	0.1804	0.2227	0.1596	----	----	0.3645	----
N4	A8	A32	A37	A46	A12	A55	A6	A38

Table 3.3: Average discrimination error (ADE) for each pairwise comparison of N4 calls selected by hand from all available N4 calls of A pod members. See legend of Table 3.2 for further details.

For the N5 call-type, only four comparisons were possible between the calls of individuals from a single matrilineal unit: A32, A46 and A37 (Table 3.4). Ranked by comparison type, mean ADE occurred in the following order from least to greatest: within MU comparison, self comparisons (Table 3.7).

A8	----							
A32	----	----						
A37	----	0.3838	----					
A46	----	0.4879	0.4067	0.5381				
A12	----	----	----	----	----			
A55	----	----	----	----	----	----		
A6	----	----	----	----	----	----	----	
A38	----	----	----	----	----	----	----	----
N5	A8	A32	A37	A46	A12	A55	A6	A38

Table 3.4: Average discrimination error (ADE) for each pairwise comparison of N5 calls selected by hand from all available N5 calls of A pod members. See legend of Table 3.2 for further details.

For call type N9, comparisons were possible between the calls of three individuals from two matriline: A32, A38, and A46 (Table 3.5). The three comparison types can be ranked in the following order from least to greatest mean ADE; between MU, within MU, and self comparisons (Table 3.7).

A8	----							
A32	0.0975	0.6472						
A37	----	----	----					
A46	0.1020	0.3722	----	0.5653				
A12	----	----	----	----	----			
A55	----	----	----	----	----	----		
A6	----	----	----	----	----	----	----	
A38	0.0884	0.1952	----	0.2026	----	----	----	----
N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 3.5: Average discrimination error (ADE) for each pairwise comparison of N9 calls selected by hand from all available N9 calls of A pod members. See legend of Table 3.2 for further details.

Hand Selection Method – W pod

Only one comparison was possible for the three W pod individuals. Since there is only one matriline within this pod, no between MU comparisons were possible and only one individual had more than 15 calls (Table 3.6).

W2	0.5540		
W3	----	----	
W5	----	----	----
N33	W2	W3	W5

Table 3.6: Average discrimination error (ADE) for each pairwise comparison of N33 calls selected by hand from all available N33 calls from W pod members. See legend of Table 3.2 for further details.

Due to the minimum input size of 15 calls per individual, hand selected comparisons were only possible for four A-pod call types (N2, N4, N5 and N9), and one W-pod call type (N33) producing a total of thirty-five comparisons (18 between-MU, 9 within-MU and 8 self) (Table 3.7).

Comparison Type	N2	N4	N5	N9	N33	All CTs
Btw MU						
N	2	11	----	5	----	18
Mean (SD)	0.233 (0.028)	0.198 (0.065)	----	0.137 (0.046)	----	0.185 (0.066)
W (p)	----	0.930 (0.412)	----	0.771 (0.046)	----	
Within MU						
N	1	4	3	1	----	9
Mean (SD)	0.317	0.385 (0.037)	0.426 (0.055)	0.372	----	0.390 (0.050)
W (p)	----	0.675 (0.006)	0.905 (0.403)	----	----	
Self						
N	2	2	1	2	1	8
Mean (SD)	0.555 (0.024)	0.555 (0.002)	0.538	0.606 (0.058)	0.554	0.566 (0.035)
W (p)	----	----	----	----	----	
F (p)	1.417 (0.890)	2.564 (0.113)	----	0.564 (0.580)	----	

Table 3.7: Summary of mean discrimination error by call type and comparison type for networks run with hand selected input sets of 15 calls. Values represent sample size (N), mean, standard deviation (SD), and results of the Shapiro-Wilks (W, p) and Levene's (F, p) tests. See legend of Table 3.2 for further details.

Hand Selection Method – Summary

For comparisons performed with hand selected input sets, the interaction between comparison type and call type was not statistically significantly different ($F_{7,23}=1.34, p=0.28$). When compared independently, the three comparison types did demonstrate significantly different levels of similarity ($F_{2,23}=33.10, p<0.05$). Similarity levels did not differ significantly between call types ($F_{2,23}=0.07, p=0.93$), as Figure 3.2 shows.

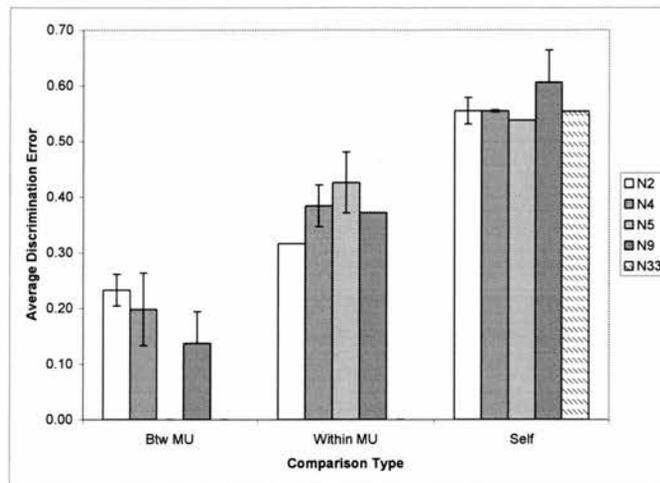


Figure 3.2: Discrimination error averaged by comparison type and call type for all hand selected comparisons performed with 15 selected calls per individual. Bars represent results for each separate call-type averaged for all individual comparisons. Error bars represent standard deviation.

Random Selection Method – A pod

Only three comparisons were possible for call type N1 and these were between calls from the three individuals of the A36 matriline; A32, A37 and A46 (Table 3.8). No ranking was possible as only comparisons of one type were performed. Table 3.14 lists these results, averaged by call type and comparison type.

A8	----							
A32	----	----						
A37	----	0.4854	----					
A46	----	0.5061	0.3800	----				
A12	----	----	----	----	----			
A55	----	----	----	----	----	----		
A6	----	----	----	----	----	----	----	
A38	----	----	----	----	----	----	----	----
N1	A8	A32	A37	A46	A12	A55	A6	A38

Table 3.8: Average discrimination error (ADE) for each pairwise comparison of N1 calls selected randomly from all available N1 calls of A pod individuals. Individuals being compared are listed vertically and horizontally. Values generated from averaging thirty trials of input sets of 8 calls per individual (16 calls per comparison). Normal text/clear box represents comparisons including members from different matrilineal units, bold text/light gray box represents comparisons within a matrilineal unit, and italic text/dark gray box corresponds to comparisons of calls produced by the same animal. Dashed lines indicate sample size was too small to perform a comparison.

For the N2 call type, a total of ten comparisons were possible between four individuals of three matriline: A8, A12, A32, and A46 (Table 3.9). The three comparison types can be ranked in the following order from least to greatest mean ADE: between MU, within MU and self comparisons (Table 3.14).

A8	0.5777								
A32	0.2619	0.6871							
A37	----	----	----						
A46	0.3092	0.3308	----	0.6809					
A12	0.2814	0.2010	----	0.1949	----				
A55	----	----	----	----	----	----			
A6	----	----	----	----	----	----	----		
A38	----	----	----	----	----	----	----	----	
N2	A8	A32	A37	A46	A12	A55	A6	A38	

Table 3.9: Average discrimination error (ADE) for each pairwise comparison of N2 calls randomly selected from all available N2 calls of A pod individuals. See legend of Table 3.8 for further details.

For the N4 call type, there were twenty-seven possible comparisons from the following seven individuals of four different matriline: A6, A8, A32, A37, A38, A46, and A55 (Table 3.10). The three comparison types can be ranked in the following order from least to greatest mean ADE: between MU, within MU and self comparisons (Table 3.14).

A8	0.5678								
A32	0.2452	0.6249							
A37	0.2710	0.3729	0.5933						
A46	0.1667	0.4126	0.3599	0.5664					
A12	----	----	----	----	----				
A55	0.0957	0.2095	0.2132	0.3301	----	----			
A6	0.1216	0.2062	0.2503	0.2465	----	0.3633	0.6123		
A38	0.0529	0.1528	0.1137	0.1682	----	0.2548	0.4989	0.5764	
N4	A8	A32	A37	A46	A12	A55	A6	A38	

Table 3.10: Average discrimination error (ADE) for each pairwise comparison of N4 calls randomly selected from all available N4 calls of A pod individuals. See legend of Table 3.8 for further details.

For the N5 call-type, six comparisons were made between A32, A37 and A46, the three individuals of the A36 matriline (Table 3.11). Ranking these

comparisons by type resulted in mean ADE in the following order from least to greatest; between MU, within MU and self comparisons (Table 3.14).

A8	----							
A32	----	<i>0.6206</i>						
A37	----	0.3497	<i>0.6108</i>					
A46	----	0.5439	0.4189	<i>0.5609</i>				
A12	----	----	----	----	----			
A55	----	----	----	----	----	----		
A6	----	----	----	----	----	----	----	
A38	----	----	----	----	----	----	----	----
N5	A8	A32	A37	A46	A12	A55	A6	A38

Table 3.11: Average discrimination error (ADE) for each pairwise comparison of N5 calls randomly selected from all available N5 calls of A pod individuals. See legend of Table 3.8 for further details.

For the N9 call type twenty-eight comparisons were possible with calls from individuals of four matriline: A8, A32, A37, A38, A46, and A55 (Table 3.12). The three comparison types can be ranked in the following order from least to greatest mean ADE: between MU, within MU and self comparisons (Table 3.14).

A8	<i>0.6300</i>							
A32	0.1664	<i>0.6323</i>						
A37	0.1360	0.3718	----					
A46	0.1851	0.3525	0.4388	<i>0.5518</i>				
A12	----	----	----	----	----			
A55	0.1672	0.2054	0.1579	0.3215	----	----		
A6	----	----	----	----	----	----	----	
A38	0.1169	0.1303	0.1099	0.2340	----	0.3062	----	<i>0.6033</i>
N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 3.12: Average discrimination error (ADE) for each pairwise comparison of N9 calls randomly selected from all available N9 calls of A pod individuals. See legend of Table 3.8 for further details.

Random Selection Method – W pod

Since only one comparison was possible within the N32 call-type, it was not included in further analysis. For the N33 call-type there were four comparisons possible from individuals of the only matriline in W pod: W2, W3 and W5 (Table 3.13). The two comparison types possible for W individuals can be ranked in the following order from least to greatest mean ADE: within MU then self comparisons (Table 3.14).

W2	----			
W3	0.3272	----		
W5	----	----	----	
N32	W2	W3	W5	

W2	0.5284			
W3	0.4767	----		
W5	0.5403	0.4196	----	
N33	W2	W3	W5	

Table 3.13: Average discrimination error (ADE) for each pairwise comparison of N32 and N33 calls randomly selected from all available N32 and N33 calls of W pod individuals. Comparisons of each call type were performed independently. See legend of Table 3.8 for further details.

Random Selection Method – Summary

Overall, randomly selecting calls for input sets made many more comparisons possible because it required fewer calls to perform a comparison. For the five most common A-pod call-types (N1, N2, N4, N5 and N9) a total of seventy-one comparisons (34 between MU, 17 within MU and 20 self) were possible instead of only thirty-five. Between the two W-pod call-types (N32 and N33), five comparisons were made (4 within-MU and 1 self) (Table 3.14). For the following analysis, the comparison results of call-types N1 and N32 were excluded as comparisons were possible for only one comparison type (Table 3.14).

A total of seven networks were run on control sets, with at least one control for each call type (N2, N4, N5 and N9) and for three individuals (A8, A37

and A46). These networks had a mean discrimination error of 0.716 (± 0.035 SD), but as they did not satisfy the ANOVA requirement for equality of variance according to Levene's test ($F=3.180, p=0.03$), a non-parametric Kruskal-Wallis test was used to demonstrate that the error generated by chance discrimination was significantly different from those discrimination errors measured for various levels of social affiliation ($H=63.438, \chi^2=3, p<0.05$).

Comparison Type	N1	N2	N4	N5	N9	N33	All CTs
Btw MU	N	5	17	---	12	---	34
	Mean (SD)	0.250 (0.050)	0.204 (0.082)	---	0.186 (0.070)	---	0.204 (0.075)
	W (p)	0.903 (0.429)	0.981 (0.968)	---	0.883 (0.095)	---	---
Within MU	N	1	4	3	3	3	17
	Mean (SD)	0.457 (0.068)	0.331	0.411 (0.063)	0.438 (0.098)	0.388 (0.045)	0.427 (0.070)
	W (p)	0.869 (0.293)	---	0.885 (0.359)	0.973 (0.686)	0.908 (0.410)	0.999 (0.941)
Self	N	4	6	3	6	1	20
	Mean (SD)	---	0.678 (0.077)	0.590 (0.024)	0.597 (0.032)	0.629 (0.049)	0.617 (0.058)
	W (p)	---	0.792 (0.096)	0.899 (0.369)	0.869 (0.293)	0.850 (0.227)	---
	F (p)	---	0.667 (0.653)	2.425 (0.110)	1.305 (0.317)	0.564 (0.580)	---

Table 3.14: Summary of mean discrimination error by call type and comparison type for networks run with randomly selected input sets of 8 calls per individual. Values show sample size (N), mean, standard deviation (SD) and results of Shapiro-Wilks (W, p) and Levene's (F, p) tests. See legend of Table 3.8 for further details.

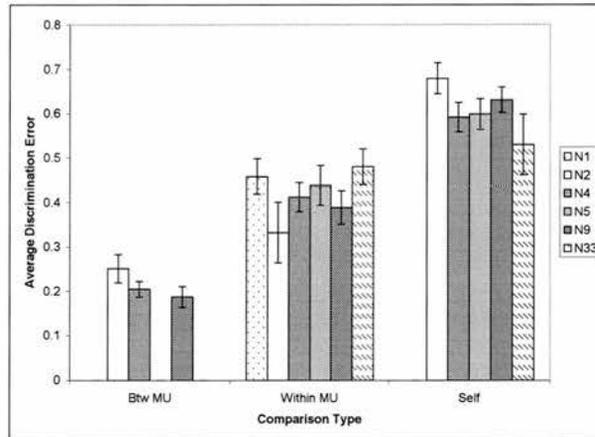


Figure 3.3: Discrimination error for all pairwise comparisons performed through random selection of eight calls per individual (16 calls per comparison). Bars represent results for each call-type averaged for all individual comparisons. Error bars represent standard deviation.

Comparisons made with randomly selected calls produced three statistically significantly different levels of acoustic similarity when grouped by the type of comparison ($F_{2,52}=37.70, p<0.006$). Similarity levels between

different call types was not significantly different ($F_{3,52}=0.35, p=0.79$). There was not a significant interaction between call type and comparison type ($F_{7,52}=1.17, p=0.34$) (Figure 3.3). Post-hoc tests confirmed that all three comparison types differed significantly (between MU/within MU: $x=0.22, p<0.05$; between MU/self: $x=0.40, p<0.05$; within MU/self: $x=0.18, p<0.05$).

Comparison of Selection Methods

Based on the results of a three-way ANOVA, network performance did not differ significantly between the two methods explored in this chapter ($F_{1,75}=0.87, p=0.36$). The methods of call selection did not interact significantly with either comparison type ($F_{2,75}=0.05, p=0.95$) or call type ($F_{2,75}=0.10, p=0.91$) and there was not a significant interaction between all three factors ($F_{7,75}=0.31, p=0.95$). Therefore, selection method did not influence the similarity levels measured by the network (Figure 3.4).

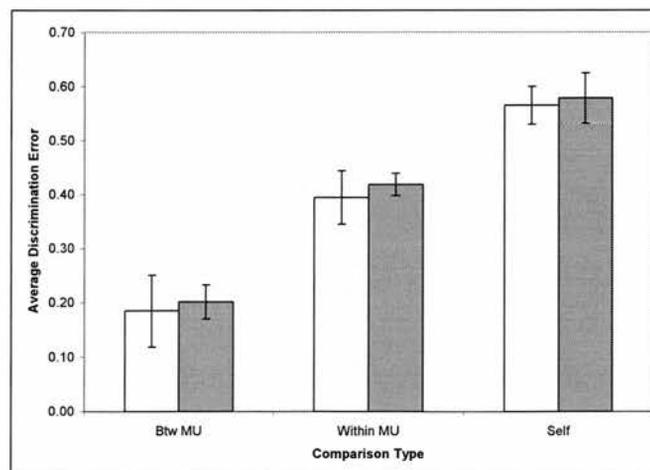


Figure 3.4: Comparison of network results from hand-selected and randomly selected input sets averaged by comparison type across all call types. Open bars represent results of hand-selected comparisons, filled bars represent results of randomly selected comparisons. Error bars represent standard error. See Tables 3.8 and 3.14 for mean values of this diagram.

DISCUSSION

Neural network analysis of the low-frequency component (LFC) contour of call types shared by different killer whales demonstrated three levels of acoustic similarity within the vocalizations of A and W pod members. Distinctly different calls have features that are easily distinguishable and therefore a lower discrimination error while calls with very similar features would result in more incorrect network classifications and a higher discrimination error. A lower amount of social affiliation between individuals was correlated with a lower discrimination error indicating the contour features of calls from more distant relatives are increasingly less similar.

Comparisons performed at the lowest level of social affiliation (between matrilineal units) had the lowest discrimination value and confirmed the previous studies of vocal variation between matrilineal units (Ford, 1991; Miller & Bain, 2000; Deecke *et al.*, 1999; Deecke, 2000) as the networks made fewer errors in discriminating between calls from separate matrilineal units. Within MU variation was measured by the network, although to a lesser extent. Finally, the comparison of different calls from the same individual demonstrated the level of discrimination error caused by intra-individual variation. The pattern of increasing vocal similarity is present in the three A pod call types for which comparisons were possible for all three comparison types (N2, N4, and N9). Additionally comparisons were possible for two comparison types for call types N1 and N5 and the level of similarity between comparison types was also

significantly different. These results suggested that levels of similarity were ordered regardless of call type.

The high level of similarity measured for self comparisons reflected the difficulty in discriminating between the two patterns of this comparison type, and demonstrated close to the minimum level of discrimination possible in network analysis, as shown by the similarity between self and control comparison levels. As the discrimination error between self and control comparisons was significantly different, vocal similarity at the level of the individual is likely influenced by either the immediate social affiliates or motivational state, both which vary between recording sessions. The fact that both of these levels are greater than 0.50 is likely because discrimination error is not only a calculation of correct to incorrect classifications, but instead consists of a more qualitative value.

Calls produced by an individual do vary, possibly stemming from the situation in which the calls were produced. The presence of conspecifics within acoustic range has been shown to increase the level of vocal activity in killer whales (Miller *et al.*, 2004; Ford, 1991; Morton *et al.*, 1986). Behavioral activities also influence the amount and type of vocalizations produced, with foraging and milling behaviors associated with the highest number of vocalizations per minute (Hoelzel & Osborne, 1986; Morton *et al.*, 1986; Ford, 1989). Vocalizations produced at different depths would be subject to varying amounts of pressure, causing changes to the shape of the air-sacs involved in sound production and therefore alterations in a single animal's voice

characteristics while diving (Tyack, 1998; Ridgeway *et al.*, 1997). Recordings were not screened for any of these criteria so, in order to reduce the effect of recording context, all available recording sessions were used in an effort to include the widest variety of emotional states and activity levels.

According to the network results, closely affiliated individuals produced calls that were less similar to each other than those from a single individual. However, this variation does not immediately suggest individual signatures. Voice characteristics may be unique to each animal simply due to differences in the physical structure of their vocal tracts. Due to the focal follow technique, calls were usually only ascribable to a single focal animal during each follow. Therefore, calls of different individuals most often come from different recording sessions, inherently including vocalizations from different contexts for each individual. The relationship or identity of individuals within a matrilineal unit may also influence vocal similarity. For example, a large portion of this dataset is based on the vocalizations of the three A36 males. As these three animals are siblings, they could demonstrate higher vocal similarity than more distantly related individuals of the same matrilineal unit, or vocal similarity that is caused by vocal differences related to their sex.

Finally, individuals have been shown to modify their vocalizations in response to other situational factors. For example, Foote *et al.* (2004) demonstrated that call duration is lengthened in the presence of boat noise. Under such circumstances the stereotyped components of killer whale calls, which define the different call types, may not be altered statically between calls so the

duration of components may not be consistent within calls from the same individual. Some components, such as initial upsweeps and terminal notes, are much more susceptible to background noise and therefore appear less consistently within the frequency contour. None of these temporary alterations were accounted for in comparison between calls but might be causes of intra-individual variation that are not individual signatures. Further studies are necessary to exclude the influence of the variables listed above on individual vocal similarity. In spite of these possible sources of variation, the network was able to discriminate between the calls of individuals from the same matrilineal group more than would be expected by chance classification. Calls of related individuals are not exactly the same, and have some distinct feature(s) although these features are not as clearly distinguishable as those between matrilineal groups.

Even though individual variation was not isolated during this experiment and group signatures are much more easily recognizable, this experiment suggests that there are individually specific features between family members. The comparisons performed in this experiment also tested two methods to find the best way to select calls for network input. The results demonstrated that the method of call selection did not significantly influence the level of similarity that was measured. Similar trends appeared regardless of whether the input sets were randomly selected from the overall dataset or hand selected according to context. Although discrimination error was consistently higher for the random selection method, the difference was well within the deviation and probably reflects the

inclusion of more intra-individual variation from context. However, the deciding factor in choosing a selection method was the much larger number of comparisons possible using the random selection method. This procedure increased the number of comparisons from thirty-three comparisons (hand selected input sets) to seventy-two, making it more useful as it incorporated all available recordings into the neural network analysis.

CHAPTER 4:
*The evaluation of temporal influence on acoustic
similarity between individuals with different social
relationships*

INTRODUCTION

Killer whales have been shown to live in stable, long-term groups, demonstrating the importance of group-living within this species (Baird, 2000; Ford *et al.*, 2000). Social interaction of this extent makes studying single or isolated vocalizations ineffective in understanding the selection pressures acting on these animals in the wild. Ignoring the presence of other animals on an individual's calls neglects an essential feature of their natural environment. The comparisons described in the previous chapter did not account for the potential influence that social interactions may place on the vocalizations of individuals. The experiment described in this chapter investigated the possibility that inter-individual interactions as demonstrated by calling bouts influence the level of vocal similarity. Using the same neural network protocol as in the previous chapter, the level of similarity found between calls selected entirely at random was compared with the level of similarity found between calls from distinctly different calling sequences.

The occurrence of a group of calls cannot be assumed to be an interaction between individuals as the vocalizations within that group are not necessarily independent. If the calls originated from two or more separate individuals in what is known as a 'group bout', these animals could be responding to each other's

calls or could be responding to similar environmental or situational factors

(Figure 4.1).

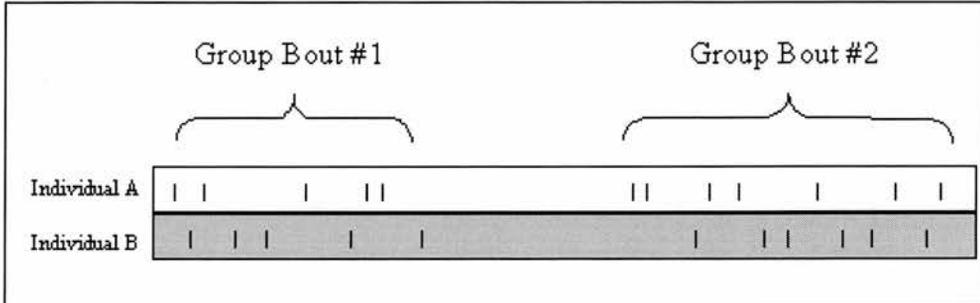


Figure 4.1: Diagram describing two group calling bouts that involve two individuals. Vertical lines represent single calls from each individual, upper white row indicates calling sequence of one individual, and lower grey row indicates calling sequence of second individual.

If one animal is producing multiple calls in a close temporal interval, also called an ‘individual bout’, the group of calls could simply be a specific type of vocalization or again, an extended response to a single stimulus. These individual bouts could be part of a larger group bout, or could occur irrespective to the presence of other individuals (Figure 4.2).

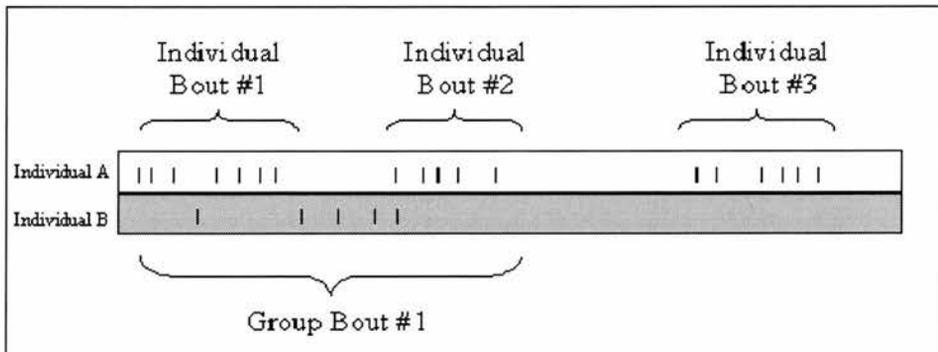


Figure 4.2: Diagram describing a group calling bout between two individuals in which the first individual produces three individual calling bouts. Vertical lines represent single calls from each individual, upper white box shows sequence of calls made by first individual and lower grey box shows sequence of calls from second individual.

In the situation of a group bout, calls produced within a group sequence could be unrelated to the calls of other individuals and should not be assumed to indicate

an interaction (Slater, 1973). Calling sequences of red jungle fowl (*Gallus gallus*) demonstrate two considerations important when analyzing calls that come from calling bouts. One rooster will crow multiple times in response to dawn, providing an example of an individual call bout in which calls become increasingly more varied through the duration of the sequence (Koene, 1996). If multiple roosters call during the same sunrise this situation would be described as a group calling bout where the individuals are not necessarily interacting (although they could be).

The extraction of calls to describe an individual's typical vocal behavior can be influenced by both individual and group bouts. Within bouts of a single individual, some vocalizations contain repeated components or are used in a particular order, so selecting calls randomly from the entire sequence would not be representative of the overall calling behavior. For example, the components, or 'themes' of humpback whale (*Megaptera novaengliae*) song occur in a specific order and each theme consists of a series of repeated sounds (Payne & McVay, 1971). Calls selected from the same theme would be much more similar than calls selected from different themes, even though both are contained within the same song. Similarly, song types of the Northern cardinal, *Cardinalis cardinalis*, are not produced at random, so a second song type will preferentially follow certain specific vocalizations (Lemon & Chatfield, 1971). These examples show that all calls from a single calling sequence are not completely independent events and randomly selecting calls from dependent sequences may influence the amount of variation measured.

The vocal behavior of resident killer whales is more likely to be an example of group bouts, where the call of one animal incites a second to call and so on (Slater, 1973). Northern resident killer whales have been shown to vocalize in group bouts, where animals call more frequently after the call of another and are more likely to match call type (Miller *et al.*, 2004). Studies of bottlenose dolphins (*Tursiops truncatus*) have documented that the calls within group bouts are not completely independent, as interacting bottlenose dolphins will not only match whistle type (Janik, 2000), but will use more structurally similar signature whistles (Watwood *et al.*, 2004; Tyack, 1986; Cook *et al.*, 2004). Southern right whales (*Eubalaena australis*) are likely to call when played a sound from another right whale, suggesting that they also call in group bouts (Clark & Clark, 1980), and sperm whales produce social 'coda' vocalizations most often in close temporal association (Weilgart & Whitehead, 1993; Whitehead & Weilgart, 1991). This is not only a characteristic of cetacean behavior, as song sparrows (*Melospiza melodia*) will respond to calling neighbors with a song of the same type (Beecher *et al.*, 1996), demonstrating that the calls of one individual are not random with respect to the calls of others.

However, the exact influence of group bouts on individual vocal behavior is not well understood. For these reasons, vocalizations within the same group bout should not be assumed to be unique and isolated from the rest of the calling sequence. If killer whales are matching call type, they could also be matching the structural features within another individual's vocalization, temporarily reducing the amount of individual variation and presenting a call that is not

typical of the individual but the group interaction. The influence of calling sequence on amount of similarity should be tested to confirm that any similarity measured is not caused by the selection of multiple calls from the same group bout.

METHODS

This method of selection was used on the same dataset of northern resident killer whale calls as the previous experiment. Chapter 2 contains complete details about the recording and processing of these data.

Group calling bouts were defined as all calls separated by less than 20 seconds, regardless of the individual that produced the call. This interval was selected following the work of Miller *et al.* (2004), which documented that killer whale are four times more likely than by chance to vocalize within 19.6 seconds of another vocalization. Every call was numerically labeled as belonging to a particular group bout and then eight calls of the same call type were randomly selected from different bouts. This was done for the calls produced by two individuals that were being compared by the neural network procedure. Thus, the input sets consisted of a total of sixteen calls (eight per individual) that corresponded with sixteen target output sets for presentation to the network. Neural networks were trained using the procedure for randomly selecting calls. Thirty different input sets of sixteen calls were selected, and run through the network a total of thirty times, after which the average from all thirty trials was calculated (see Chapter 2 for random selection procedure and network

manipulation). Network output measures the amount of similarity between two calls by calculating the average discrimination error (ADE), or the number of calls classified to the correct pattern as compared to the number classified to the incorrect pattern. Therefore, more similar calls would result in a higher ADE, as they would be more difficult to distinguish.

As the results of the first experiment did not show any significant difference between the results from either selection type (calls selected by hand, or calls selected randomly by computer), both appeared to choose calls representative of an individual. Due to the low number of comparisons resulting from call sets selected by hand and therefore including a limited portion of the available data, results of the random selection procedure were used to assess the possibility that call similarity may be influenced by group calling bouts. Resulting ADE was divided into three groups based on the level of social affiliation between the individuals that produced the calls; between matrilineal unit (MU) comparisons for comparisons between individuals from separate family groups, within-MU comparisons for comparisons between two individuals from the same family group and self comparisons for comparisons of calls from the same individual. A three-way analysis of variance test was performed using SYSTAT software to compare comparison type, call type and selection method. It is possible that a Type I error originated from multiple testing, so a Bonferroni correction was used to adjust the significance level to 0.017. The Shapiro-Wilks test was also performed to confirm a normal distribution for these data and the

Levene's test was used to determine equivalence of variance between groups (Zar, 1996).

RESULTS

This analysis was conducted on the same dataset as the one included in the previous chapter. Therefore, Chapter 3 contains a detailed description of the available dataset and information about the relationship between the individual animals mentioned here (see pages 3-7 to 3-9). Table 4.1 summarizes the results of separating the dataset for this analysis.

	Max	Min	Mean	Mean # calls
Inter-Bout Interval	00:23:26:50	00:00:20:45	00:02:09:18	---
Intra-Bout Interval	00:00:19:20	00:00:00:21	00:00:06:06	4

Table 4.1: Descriptive summary of bout separations for the time interval between calls of separate bouts (inter-bout interval) and between calls within a group bout (intra-interval bout). Maximum, minimum and mean intervals all displayed in hour:minute:second:tenths format.

Random Selection

To briefly summarize the results of the previous chapter, for call types shared by group members, comparisons performed with the random selection procedure produced three levels of call similarity that correlated with the social relationships between the individuals that produced them (Chapter 3). Between MU comparisons had the lowest mean discrimination error (0.204 ± 0.075 SD) within MU comparisons had a higher mean discrimination error (0.427 ± 0.070) and self comparisons had the highest mean discrimination error (0.617 ± 0.058) for comparisons performed with randomly selected calls. The results when grouped by comparison type produced statistically significant differences ($F_{2,52}=37.70, p<0.005$). In contrast the social relationship of the individuals

producing calls did not statistically alter the discrimination for different call types ($F_{3,52}=0.35, p=0.79$).

Selection from Different Group Bouts

For the N2 call-type, the available data made it possible to make four comparisons from the calls of individuals (A8, A32 and A46) from two matriline (Table 4.2). When averaged by comparison type, average discrimination error (ADE) resulted in the following order listed from least to greatest; between-MU, within-MU and self comparisons (Table 4.5).

A8	<i>0.5033</i>							
A32	0.3484	----						
A37	----	----	----					
A46	0.3622	0.4096	----	----				
A12	----	----	----	----	----			
A55	----	----	----	----	----	----		
A6	----	----	----	----	----	----	----	
A38	----	----	----	----	----	----	----	----
N2	A8	A32	A37	A46	A12	A55	A6	A38

Table 4.2: Average discrimination error (ADE) for each pairwise comparison of N2 calls selected from group bouts. Individual whales are listed vertically and horizontally. Values generated from 8 vocalizations per individual (16 calls per comparison). Normal text/clear box represents comparisons including members from different matrilineal units, bold text/light gray box represents comparisons within a matrilineal unit, and italic text/dark gray box corresponds to comparisons of calls produced by the same animal. Dashed lines indicate sample size was too small to perform a comparison.

For the N4 call-type there were sixteen comparisons made from the calls of six individuals (A8, A32, A37, A46, A55 and A38) from four different matriline (Table 4.3). Mean ADE resulted in the following order from least to greatest; between MU, within MU and self comparisons (Table 4.5).

A8	0.5122								
A32	0.2517	----							
A37	0.338823	0.4084	----						
A46	0.1452	0.4836	0.3838	0.5296					
A12	----	----	----	----	----				
A55	0.0658	0.2367	0.1999	0.3101	----	----			
A6	----	----	----	----	----	----	----		
A38	0.1090	0.1788	0.1696	0.2613	----	----	----	----	
N4	A8	A32	A37	A46	A12	A55	A6	A38	

Table 4.3: Average discrimination error (ADE) for each pairwise comparison of N4 calls selected from group bouts. Individual whales are listed vertically and horizontally. See legend of Table 4.1 for further details.

For the N9 call type, there were six comparisons made between the calls of three individuals (A8, A32 and A46) from two matriline (Table 4.4). Mean ADE resulted in comparison types in the following order from least to greatest; between MU, within MU and self comparisons (Table 4.5).

A8	0.6815								
A32	0.1694	0.4826							
A37	----	----	----						
A46	0.2150	0.3850	----	0.5032					
A12	----	----	----	----	----				
A55	----	----	----	----	----	----			
A6	----	----	----	----	----	----	----		
A38	----	----	----	----	----	----	----	----	
N9	A8	A32	A37	A46	A12	A55	A6	A38	

Table 4.4: Average discrimination error (ADE) for each pairwise comparison of N9 calls selected from group bouts. Individual whales are listed vertically and horizontally. See legend of Table 4.1 for further details.

Bout separated comparisons of all three call-types resulted in 26 comparisons with mean discrimination error in the following order from least to greatest; between MU, within MU and self comparisons (Table 4.5). As with the randomly selected comparisons, there was no significant interaction between

comparison type and call type ($F_{4,16}=1.47, p=0.26$). Call type did not differ significantly according to statistical comparison ($F_{2,16}=0.348, P=0.71$), but similarity was significantly influenced by comparison type ($F_{2,16}=21.70, p<0.05$) (Figure 4.3).

Comparison Type		N2	N4	N9	All CTs		
Btw MU	N	2	11	2	15		
	Mean (SD)	0.355 (0.010)	0.206 (0.083)	0.192 (0.032)	0.224 (0.089)		
	W (p)	----	0.986 (0.990)	----			
Within MU	N	1	3	1	5		
	Mean (SD)	0.410	0.425 (0.052)	0.385	0.414 (0.041)		
	W (p)	----	0.921 (0.456)	----			
Self	N	1	2	3	6		
	Mean (SD)	0.503	0.521 (0.012)	0.556 (0.109)	0.535 (0.073)		
	W (p)	----	----	0.827 (0.180)			
			F (p)	----	1.888 (0.191)	0.065 (0.319)	

Table 4.5: Summary of mean discrimination error by call type and comparison type for networks run with calls selected according to bout context. Values show sample size (N), mean, standard deviation (SD) and results of Shapiro-Wilks (W, p) and Levene's (F, p) tests.

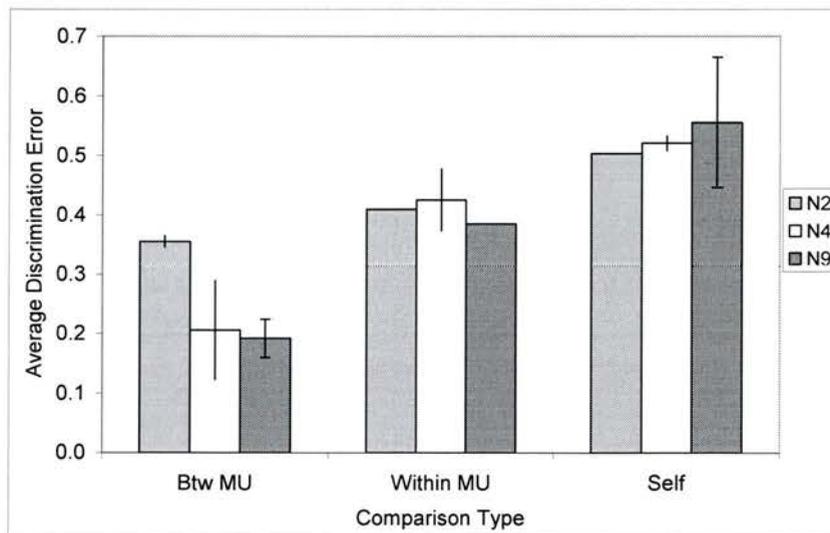


Figure 4.3: Mean similarity levels for comparisons run with calls from separate group bouts, separated by comparison type and call-types. Networks run with eight calls selected from group bouts more than 20 seconds apart. Error bars represent standard deviation.

Comparison of selection methods

There was not a significant interaction between the three factors analyzed ($F_{5,68}=0.97, p=0.44$). The results did not differ significantly based on selection method ($F_{1,68}=0.003, p=0.96$), although further comparisons with more data would be useful to strengthen this result. Additionally, selection method did not influence either comparison type ($F_{1,68}=0.41, p=0.52$), or call type ($F_{4,68}=0.22, p=0.93$). Using a single call per group bout sequence reduced the number of comparisons from seventy-two to twenty-six. This limitation resulted in comparisons being possible for only three call-types (N2, N4 and N9), whereas five call-types could be compared when calls were randomly selected (N2, N4, N5, N9 and N33). Grouped by comparison type for calls selected from different group bouts, there were fifteen between MUs, five within MUs and six self comparisons, all between A pod individuals. No comparisons were possible for W-pod individuals. With a much larger dataset to select from, randomly selected comparisons allowed seventy-one comparisons (34 between MU, 17 within MU and 20 self comparisons) and therefore this selection method was chosen as a more functional analysis method. Figure 3.2 contains ADE grouped by comparison type and Table 3.14 lists the mean values for these comparisons.

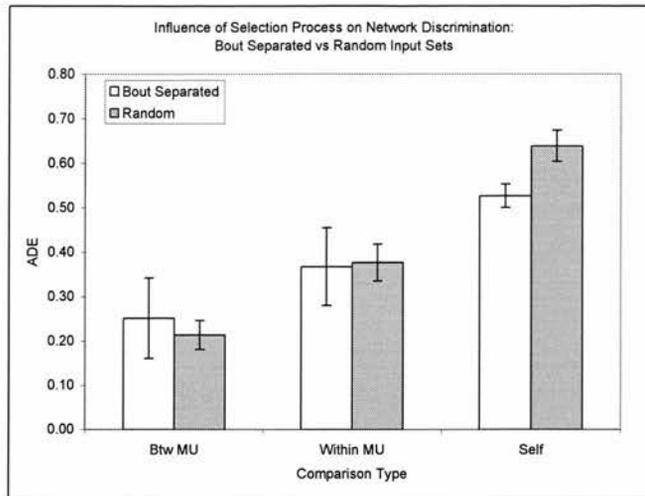


Figure 4.4: Comparison between calls selected entirely randomly and from separate calling bouts. Bars with stripes represent entirely random results, while solid bars are results of separate calling bouts for each of three call types: N2, N4, and N9. Error bars represent one standard deviation.

DISCUSSION

Stereotyped calling in killer whales occurs in distinct calling bouts, where the call of one individual influences the vocal behavior of its group mates (Miller *et al.*, 2004). Direct responses to the vocal signals of a conspecific could indicate to the original caller that the signal was received. Another suggestion for the importance of this calling behavior would be to provide information about the distance, orientation and direction of travel between individuals (Miller, 2002). Other individuals also match call type within vocal exchanges (Miller *et al.*, 2004), presenting the possibility that calls are structurally modified to match each other as well. On a brief temporal scale, matching structural call features would strongly confirm a response to the signaler (Miller *et al.*, 2004). Comparing calls within the same call type from one individual raises the likelihood that two calls from the same group bout could be compared. Individual variation might be

reduced during that group bout and result in a level of similarity that is not representative of that individual's vocal communication.

Although bout criterion intervals were not calculated for each individual follow (Slater & Lester, 1982), calls used in this analysis were produced in group bouts separated by a time interval greater than 20 seconds in order to ensure there was no influence between calls included when testing their similarity. The network results still produced three levels of similarity that corresponded with the amount of social affiliation between individuals. Intra-individual calls were the most similar, intra-group calls were less similar and inter-group comparisons were the most different. Results were significantly different between these comparison types, indicating that the network was able to discriminate between individuals and suggesting that there are consistently identifiable individual features within these calls. However, bout separated selection reduced the number of comparisons by more than fifty percent. To ensure that the patterns observed are representative of an individual's behavior, the widest possible range of calls should be used, and therefore selecting calls from the entire dataset will be a more robust method of selection.

These results do not exclude the possibility that vocal matching occurs within killer whale group bouts. In this experiment, comparisons were performed to measure the similarity between calls that were carefully selected from different group bouts across calling contexts, but were never compared with calls from the same group bouts. In order to study the effects of other calling individuals, calls from all of the individuals vocalizing during a group bout should be compared for

similarity. Additionally, vocalizations must be selected for each individual according to the presence and absence of other individuals, a criterion that was not possible during this study since the identities of surrounding individuals were not usually recorded in the field. A different method of quantifying similarity as well as more efficient recording techniques of individuals will be needed to determine if any vocal matching occurs within wild killer whale groups. The focal follow technique concentrates on recording the position and behavioral observations of a single focal individual so a method to track the movement of all individuals within the area would be necessary to fully investigate the possibility of vocal exchanges. Finally, further analysis of bouting behavior would require the calculation of bout criterion intervals specific to each group of animals included, a consideration that was beyond the scope of this study. In lieu of a measured interval, the previously documented value of approximately 20 second was used, and resulted in an average inter-bout interval that was noticeably larger than the intra-bout interval. Nevertheless, these results were useful in determining that bouting context does not need to be accounted for at this level of analysis, although would be an interesting subject of further study.

CHAPTER 5:
***Northern resident vocal similarity quantified through
comparisons of separate components within and
between call types***

INTRODUCTION

The goal of this chapter is to further investigate more potential indicators of individual vocal signatures in killer whale vocalizations, namely in the other frequency component of their vocalizations, as well as understand the extent to which similarities appear with the vocal repertoire of these animals. The physical structure of killer whale vocalizations is relatively unique amongst mammals as it contains two simultaneous and independently modulated frequency contours. These two distinct components, a low frequency component (LFC) ranging from 80 to 240 Hz and a high frequency component (HFC) with a fundamental frequency between 2 and 12 kHz, each have distinctive frequency modulations (Bain 1986; Ford 1987; Miller and Bain 2000) (Figure 5.1).

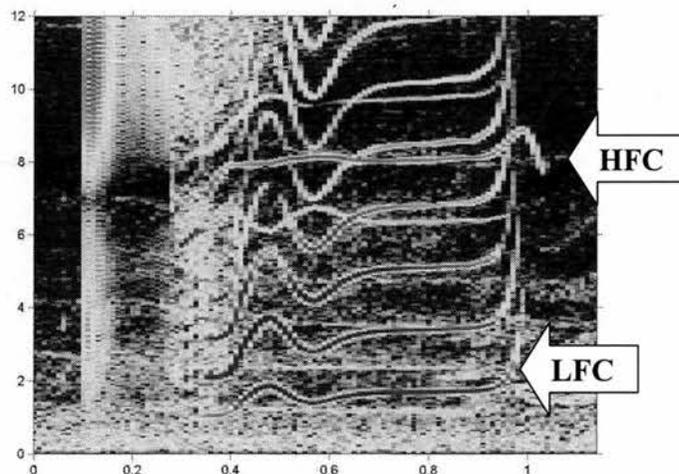


Figure 5.1: FFT spectrogram of N2 call from individual A8 showing the relationship between the two simultaneous frequency contours. The spectrogram was created with custom MATLAB software, a sampling rate of 48 kHz, an FFT size of 2048 samples, a dynamic range of 50 dB and a maximum frequency limit of 11 kHz. See Appendix I for spectrograms of other call types.

Two simultaneous vocalizations with independent frequency modulations have been described in penguins of the genus *Aptenodytes*. Through both playback studies and spectral analysis, the two-voice structure of *Aptenodytes* vocalizations has been found to be significant in the individual recognition system used by king and emperor penguin chicks to locate their parents (Robisson 1992; Robisson *et al.* 1993; Jouventin *et al.* 1999; Aubin *et al.* 2000; Lengagne *et al.* 2001). While penguins are very different to killer whales in many respects, they are forced to deal with similar ecological constraints such as how to reunite with specific individuals in a low visibility, high background noise environment with few fixed and identifiable sites. Unlike colonial penguins, nesting penguin species do not use two-voiced vocalizations, and individual signatures have not yet been found in these species (Aubin *et al.* 2000; Searby *et al.* 2004).

The exact function of each frequency component is not known. It has been suggested that killer whales may use the HFC to convey information on direction of movement to animals traveling in close proximity (Miller 2002). The highly directional nature of this component relative to the low-frequency component makes it unlikely to function in coordinating more distant animals. A second suggestion is that high-frequency sounds function in inter-pod communication, as it has been observed that call types with a HFC are used more frequently in associations where multiple pods are present than call types with only a LFC (Foote, personal comm.). Attenuation will have less influence on communication during multipod associations as all individuals are grouped within a small area. The suggestion that HFCs have fewer features that distinguish

between matrilineal groups than LFCs also suggests that this component has been selected for consistency (Miller and Bain 2000). LFC variations between different matrilineal units and pods are most likely caused by individuals not hearing the version of another matrilineal unit. If the amount of difference between versions of the same call type are indicators of how long groups have been separated (Ford 1991; Deecke 2000; Yurk *et al.* 2002), consistency suggests the opposite; that individuals are hearing and using the same version when interacting. Regardless of its exact function, the development of this unique 'two-voice' feature suggests it plays a highly significant role within the interactions of killer whales and should be considered when assessing the possibility for individual signatures. The similarity of both the LFC and HFC produced by different individual whales must be measured to more fully explore the potential for individually distinctive vocalizations to exist in resident killer whale vocalizations.

Killer whale calls have been catalogued into separate 'call types' by identifying the stereotyped series of components divided by changes in pulse-repetition rate and both components contribute to this classification as both are present in some call types, while only the LFC is present in others. The differences between call types are clear enough that calls can be classified by ear and/or visual inspection of sonagrams (Ford 1987). Resident killer whales off Vancouver Island are divided into four distinct clans based on the set of call types that they use (Ford 1991; Ford *et al.* 2000). Ford *et al.* (2000) suggest that these four clans are descended from four separate ancestral lineages, which explains the

lack of shared call types between clans existing today. Although separate pods share call types, they will use only a portion of the call types of the entire clan, and this portion will be described hereafter as the pod's 'repertoire'. Separate matrilineal units (MUs), which consist of a female and all of her offspring, use the same repertoire, although consistent variations in call types have been documented between different matrilineal units of the same pod. The comparisons performed in the previous experiments demonstrated that more distant associations resulted in less similar low-frequency components, so comparisons of calls from different clans should be measurably distinct. In this experiment that trend was extended to assess the similarity of each frequency component independently and to determine the amount of vocal similarity between individuals with the lowest amount of social affiliation.

In the concluding analysis described here, three sets of comparisons are performed to thoroughly quantify the overall influence of social affiliation on vocal similarity in the stereotyped calls of killer whales. The first set of comparisons extends the analysis described in Chapter 3 to measure similarity levels of both frequency components at different levels of association. Next, the extent of vocal similarity between call types is measured as individuals that associate infrequently do not share any call types. Call type similarity is measured both between call types that are used by affiliated individuals (like those of a matrilineal unit) and between call types of less affiliated individuals (those of different clans). Within this second set of comparisons, termed 'within shared repertoire' comparisons, similarity was measured between calls of two

different call-types produced by individuals at the same three levels of social affiliation used in the first set. Then these results were compared with those of the third set of comparisons, 'within unshared repertoire' comparisons, to determine that the measured level of vocal similarity was due to the low level of social affiliation and not the structural differences evident between different call types. These comparisons included calls of two different call types from individuals at only one level of social affiliation, termed 'between clan'. Unlike the comparisons made in previous chapters, all three sets of comparisons were performed independently on both the low-frequency component (LFC) and the high-frequency component (LFC) of available vocalizations.

METHODS

This analysis was performed on the same collection of northern resident killer whale calls as used in the two previous experiments (See chapter 2 for further details of data collection). Since these experiments demonstrated that randomly selecting from all calls recorded from an individual reduced variation from situational context and is not measurably influenced within the analysis by the calling behavior of other individuals, MATLAB software has been developed to randomly select contours from all available calls of a specific individual.

Field recordings were only available of calls from two pods; A pod, which uses the call types of A clan, and W pod, which uses the call types of R clan. Calls are available from A pod individuals of four different matriline, the details of which are contained in Table 5.1. As all W pod individuals are from the same

matriline, their genealogical grouping does not require tabulation. For this study, all call types are labeled with the letter ‘N’ and a number; single digits describe call types used by A pod and double digits describe call types used by W pod. Only a portion of each repertoire was included in this analysis as sufficient sample sets of all call types were not available. Additionally, because the similarity of each component was compared separately, only call types containing both components were used. Call types N2, N4, N5, and N9 were used to represent the repertoire of A clan, and call types N32 and N33 were used to represent the repertoire of R clan (see Appendix I for spectrograms of all analyzed call types).

Matriline	A36	A12	A30	A11	A8	A9	A23	A25
Individuals (Females)		A12, A34, A62	A30, A50, A54	A11, A24, A35, A59, A45, A52, A56	A8, A28, A42		A23, A43	A51
Individuals (Males)	A32, A37, A46	A33, A55	A6, A38, A39	A13		A5, A9	A27, A60	A61
Juveniles		A67	A72	A64, A70, A71	A66		A69	

Table 5.1: Individuals from A pods grouped by matriline and sex. Bold font indicates individuals used in this study. Individuals are listed within categories in order of their age (beginning with the oldest) (Ford *et al.*, 2000).

Pairwise comparisons were selected based on three separate categories according to repertoire level; (1) ‘within call type’ comparisons to compare calls of the same type from different individuals, (2) ‘within shared repertoire’ comparisons to compare different call types from individuals that share all of their call types, and (3) ‘within unshared repertoire’ comparisons to compare different call types from individuals that do not share any call types. Comparisons at the ‘within call type’ and ‘within repertoire’ repertoire levels used the three levels of social affiliation described in previous experiments; (1) ‘between MU’ comparisons involving individuals from different matrilineal units, (2) ‘within

MU' comparisons using two individuals from the same matrilineal unit and (3) 'self' comparisons using two sets of different calls from the same individual. As comparisons between unshared repertoires did not fall into one of the comparison types listed above, all of these comparisons were performed on a separate level of social affiliation, termed 'between clan' comparisons. Each component of the call was analyzed independently so all of the above comparisons were performed twice, once on the LFC and once on the HFC. For all of these comparisons, sets of eight calls were randomly selected from each of the individuals and presented to a neural network for pattern recognition. The network then attempted to identify each call with one of the individuals, and calculated the proportion of incorrect classifications made.

Three-way analysis of variance tests were performed to find if differences were significant between call types, comparison types and frequency components for within call type comparisons. Significance levels were adjusted when performing multiple comparisons using the Bonferroni correction to avoid Type I errors although this correction did not alter any of the factors found to be significant. Additionally, the Shapiro-Wilks and Levene's tests were used to confirm normality and homogeneity of variance respectively. Finally, a second three-way analysis of variance test was performed to investigate any differences present between repertoire levels (Zar 1996).

RESULTS

For each of the two components, comparisons were performed on the six most commonly used call-types (N2, N4, N5, N9, N32 and N33), resulting in a total of 566 comparisons. The calls of eleven different individuals (A6, A8, A12, A32, A37, A38, A46, A55, W2, W3 and W5) were used in this analysis, including five different matriline and two different clans.

Within Call Type Analysis

Low-Frequency Component

Within Call Type – Low Frequency Component – A pod

Sixty-eight comparisons were performed using call-types N1, N2, N4, N5 and N9 from two of the following individuals: A6, A8, A12, A32, A37, A38, A46, and A55. This analysis consisted of thirty-four between MU comparisons, fourteen within MU comparisons and nineteen self comparisons (Table 3.14). See chapter 3 for a detailed description of these results.

Within Call Type – Low Frequency Component – W pod

Five comparisons made with the two W pod call-types (N32 and N33) used for this analysis. For N32, there was one within MU comparison and for N33 there were three within MU and one self comparisons. When grouped by comparison type, within MU comparisons were less similar than self comparisons (Table 3.14). See chapter 3 for a detailed description of these results.

Within Call Type – Low Frequency Component – Summary

When the discrimination error was averaged by comparison type across all call types, there was a consistent trend in the results. Between MU comparisons had the lowest mean (0.2043 ± 0.0751 SD), within MU had a slightly greater mean (0.4270 ± 0.0697) and self comparisons had the greatest mean (0.6173 ± 0.0577) (Table 3.14, Figure 3.3).

High Frequency Component

Within Call Type – High Frequency Component – A pod

The N2 call-type allowed for HFC comparisons between individuals from four different matriline; A8, A12, A32, A37 and A46. There were eight comparisons in total: five between MU, one within MU and two self (Table 5.2). When grouped by comparison type, mean ADE could be ranked in the following order from least to greatest: between MU, within MU and self comparisons (Table 5.7).

A8	0.5071							
A32	0.3237	0.5407						
A37	----	----	----					
A46	0.4288	0.3892	----	----				
A12	0.3581	0.2601	----	0.2742	----			
A55	----	----	----	----	----	----		
A6	----	----	----	----	----	----	----	
A38	----	----	----	----	----	----	----	----
N2	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.2: Average discrimination error (ADE) for each pairwise HFC comparison between N2 calls randomly selected from all available N2 calls of A pod individuals. See legend of Table 5.2 for further details.

For the N4 call type, HFC comparisons were possible for individuals from four different matriline A8, A12, A32, A37, A38, A46 and A55. Twenty-seven

comparisons were performed: seventeen between MU, four within MU and six self (Table 5.3). Mean ADE when averaged by comparison type could be ranked in the following order from least to greatest: between MU, within MU and self (Table 5.7).

A8	0.5047								
A32	0.4201	0.4982							
A37	0.4242	0.3398	0.4968						
A46	0.4366	0.4298	0.4628	0.5433					
A12	----	----	----	----	----				
A55	0.2010	0.1670	0.2797	0.2766	----	----			
A6	0.2193	0.2317	0.3374	0.2857	----	0.3693	0.5559		
A38	0.1344	0.1187	0.1832	0.1824	----	0.2829	0.4730	0.5132	
N4	A8	A32	A37	A46	A12	A55	A6	A38	

Table 5.3: Average discrimination error (ADE) for each pairwise HFC comparison between N4 calls randomly selected from all available N4 calls of A pod individuals. See Table 5.2 for further details.

For the N5 call-type comparisons were performed between individuals from a single matriline; A32, A37 and A46. Only four comparisons were possible: three within MU and one self (Table 5.4). Ranked by comparison type from least to greatest, mean ADE was in this order: within MU and self comparisons (Table 5.7).

A8	----								
A32	----	----							
A37	----	0.3043	----						
A46	----	0.5483	0.3472	0.5441					
A12	----	----	----	----	----				
A55	----	----	----	----	----	----			
A6	----	----	----	----	----	----	----		
A38	----	----	----	----	----	----	----	----	
N5	A8	A32	A37	A46	A12	A55	A6	A38	

Table 5.4: Average discrimination error (ADE) for each pairwise HFC comparison between N5 calls randomly selected from all available N5 calls of A pod individuals. See legend of Table 5.2 for further details.

For the N9 call type, HFC comparisons were possible between the calls of six different animals from four different matriline: A8, A32, A37, A38, A46 and

A55 and nineteen comparisons were possible; twelve between MU, three within MU and four self (Table 5.5). Mean ADE averaged by comparison type could be ranked in the following order from least to greatest: between MU, within MU, and self comparisons (Table 5.7).

A8	0.5209								
A32	0.2160	0.5377							
A37	0.2413	0.4120	----						
A46	0.2114	0.4238	0.4690	0.5186					
A12	----	----	----	----	----				
A55	0.1119	0.1618	0.1652	0.2700	----	----			
A6	----	----	----	----	----	----	----		
A38	0.1149	0.1273	0.0959	0.1465	----	0.2505	----	----	0.5468
N9	A8	A32	A37	A46	A12	A55	A6	A38	

Table 5.5: Average discrimination error (ADE) for each pairwise HFC comparison between N9 calls randomly selected from all available N9 calls of A pod individuals. See legend of Table 5.2 for further details.

Within Call Type – High Frequency Component – W pod

Two W pod call types were produced frequently enough to be included within the analysis, but the first of these, N32, allowed only one HFC comparison so it was not included in any further analysis of within call-type similarity. HFC comparisons of call-type N33 were possible for individuals W2, W3 and W5 and resulted in four comparisons: three within MU and one self (Table 5.6). Mean ADE averaged by comparison type could be ranked in the following order from least to greatest: within MU and self comparisons (Table 5.7).

W2					W2	0.4784		
W3	0.2453	---			W3	0.4119	---	
W5	---	---	---		W5	0.2411	0.4080	---
N32	W2	W3	W5		N33	W2	W3	W5

Table 5.6: Average discrimination error (ADE) for each pairwise HFC comparison between N32 and N33 calls randomly selected from all available N32 and N33 calls of W pod individuals. See legend of Table 5.2 for further details.

Within Call Type – High Frequency Component – Summary

Results of the HFC comparisons were similar to those of the LFC.

Individuals from two different matrilineal units had the least similar calls, as demonstrated by the lowest mean ADE (0.2443 ± 0.0990), while individuals from the same family group had less distinguishable calls (0.4069 ± 0.0746). Finally, the comparison of different calls from the same individual produced the highest mean ADE (0.5219 ± 0.0232), indicating that these calls were the most similar (Table 5.7, Figure 5.2).

Comparison Type		N2	N4	N5	N9	N33	All CTs
Between MU	N	5	17	---	12	---	34
	Mean (SD)	0.329 (0.068)	0.268 (0.102)	---	0.176 (0.060)	---	0.244 (0.099)
	W (p)	0.943 (0.686)	0.939 (0.307)	---	0.930 (0.381)	---	
Within MU	N	1	4	3	3	3	14
	Mean (SD)	0.389	0.426 (0.061)	0.412 (0.118)	0.435 (0.030)	0.354 (0.098)	0.407 (0.075)
	W (p)	---	0.855 (0.244)	0.753 (0.007)	0.897 (0.377)	0.767 (0.038)	
Self	N	2	6	1	4	1	14
	Mean (SD)	0.524 (0.024)	0.519 (0.025)	0.544	0.531 (0.014)	0.4780	0.522 (0.023)
	W (p)	---	0.848 (0.153)	---	0.894 (0.400)	---	
	F (p)	8.231 (0.510)	4.075 (0.030)	---	3.481 (0.056)	---	

Table 5.7: Summary of network performance for within call type comparisons of the HFC across three comparison types. Values represent sample size (N), mean, standard deviation (SD), and results of Shapiro-Wilks (W,p) and Levene’s test (F, p) when grouped by comparison type (listed vertically) and call type (listed horizontally).

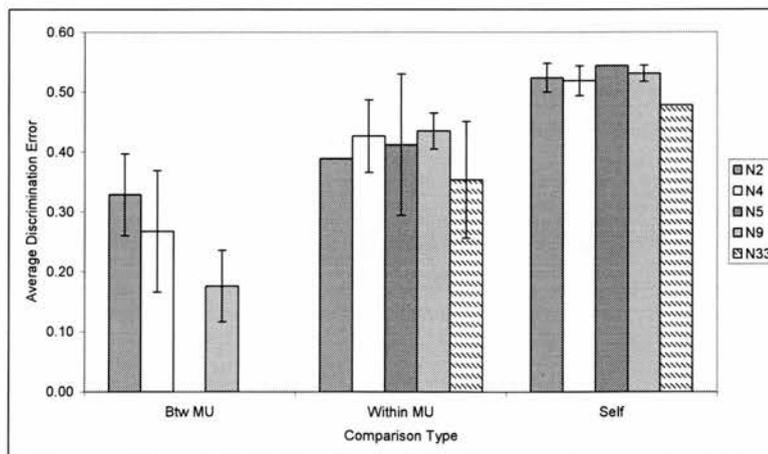


Figure 5.2: Average discrimination error (ADE) for pairwise HFC comparisons by call type across three comparison types. Comparisons performed through thirty trials containing sets of eight calls randomly selected from the entire dataset from each of two individuals (16 calls per comparison). Error bars represent standard deviation.

Within Call Type – Summary

There was no significant interaction effect between the three factors when considered together ($F_{6,49}=1.42, p=0.23$) nor when considered in pairs (call type/comparison type: $F_{3,70}=0.10, p=0.96$; call type/component: $F_{2,70}=0.506, p=0.61$; comparison type/component: $F_{1,70}=0.01, p=0.98$). The type of comparison had a significant effect on similarity ($F_{2,70}=4.58, p<0.05$) while call type ($F_{2,70}=0.333, p=0.72$) and component did not ($F_{1,70}=0.01, p=0.92$). Post hoc Tukey HSD tests confirmed that all three comparison types were significantly different (between MU/within MU: $x=0.16, p<0.05$; between MU/self: $x=0.28, p<0.05$; within MU/self: $x=0.11, p<0.05$).

Within Shared Repertoire Analysis

Low Frequency Component

The number of calls available from each individual determined the number of comparisons possible for each combination of call types. Calls of type N2 were selected for comparison from individuals A8, A12, A32 and A46, representing three matriline. Four matriline were represented in all comparisons of N4 call types, as enough calls were recorded from individuals A6, A8, A32, A37, A38, A46 and A55 for analysis. Only individuals A32, A37 and A46 used call type N5, so comparisons were only possible within a single matriline. The last A pod call type included in these comparisons was N9 and these calls were recorded from individuals A8, A32, A37, A38, A46 and A55, which represent

four matriline. Additionally, there were two call types from W pod individuals with enough calls to compare. Only two matriline use these call types and calls were recorded only from one of them. N32 calls from individuals W2 and W3 were compared with N33 calls from individuals W2, W3 and W5.

Within Shared Repertoire – Low Frequency Component – A Pod

The analysis of call types N2 and N4 produced a total of twenty-eight LFC comparisons. Twenty of those were between individuals of different matrilineal units, five were between individuals of the same matrilineal unit and three were between different calls from the same individual (Table 5.8). As Table 5.15 shows, mean ADE can be ranked by comparison type in the following order from least to greatest; between MU, within MU and self comparisons.

A8	0.0747	0.0896	0.0309	0.0438	---	0.0164	0.0163	0.0309
A32	0.0508	0.1102	0.0218	0.0805	---	0.0022	0.0088	0.0186
A37	---	---	---	---	---	---	---	---
A46	0.0218	0.0834	0.0288	0.0698	---	0.0014	0.0075	0.0088
A12	0.0510	0.0553	0.0136	0.0076	---	0.0018	0.0017	0.0050
A55	---	---	---	---	---	---	---	---
A6	---	---	---	---	---	---	---	---
A38	---	---	---	---	---	---	---	---
N2 - N4	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.8: Average discrimination error (ADE) for each pairwise LFC comparison between N2 and N4 calls randomly selected from all available N2 and N4 calls of A pod individuals. Individuals being compared are listed vertically and horizontally. See legend of Table 5.2 for further details.

Twelve comparisons were possible between call types N2 and N5. When grouped by comparison type, these comparisons produced six between MU comparisons, four within MU comparisons and two self comparisons (Table 5.9). When averaged by comparison type, mean ADE can be ranked in the following

order from least to greatest: self, between MU and within MU comparisons (Table 5.15).

A8	---	0.0270	0.0148	0.0355	---	---	---	---
A32	---	0.0030	0.0050	0.0490	---	---	---	---
A37	---	---	---	---	---	---	---	---
A46	---	0.0102	0.0078	0.0018	---	---	---	---
A12	---	0.0041	0.0018	0.0108	---	---	---	---
A55	---	---	---	---	---	---	---	---
A6	---	---	---	---	---	---	---	---
A38	---	---	---	---	---	---	---	---
N2 - N5	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.9: Average discrimination error (ADE) for each pairwise LFC comparison between N2 and N5 calls randomly selected from all available N2 and N5 calls of A pod individuals. Individuals being compared are listed vertically and horizontally. See legend of Table 5.2 for further details.

Twenty-four LFC comparisons were performed between call types N2 and N9: sixteen between MU, five within MU and three self (Table 5.10). Mean ADE can be ranked by comparison type in the following order from least to greatest: within MU, between MU and self comparisons (Table 5.15).

A8	0.2023	0.1591	0.1430	0.1829	---	0.1714	---	0.1795
A32	0.1368	0.2056	0.0920	0.0813	---	0.0948	---	0.0657
A37	---	---	---	---	---	---	---	---
A46	0.1651	0.1650	0.1553	0.1734	---	0.0982	---	0.0708
A12	0.1041	0.0973	0.0619	0.0470	---	0.0301	---	0.0299
A55	---	---	---	---	---	---	---	---
A6	---	---	---	---	---	---	---	---
A38	---	---	---	---	---	---	---	---
N2 - N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.10: Average discrimination error (ADE) for each pairwise LFC comparison between N2 and N9 calls randomly selected from all available N2 and N9 calls of A pod individuals. Individuals being compared are listed vertically and horizontally. See legend of Table 5.2 for further details.

Twenty-one LFC comparisons were made between call types N4 and N5: twelve between MU, six within MU and three self (Table 5.11). Mean ADE can be ranked by comparison type in the following order from least to greatest: between MU, within MU and self comparisons (Table 5.15).

A8	—	0.0800	0.1174	0.0758	—	—	—	—
A32	—	<i>0.1366</i>	0.0700	0.0636	—	—	—	—
A37	—	0.1333	<i>0.1483</i>	0.1349	—	—	—	—
A46	—	0.0410	0.0569	<i>0.0315</i>	—	—	—	—
A12	—	—	—	—	—	—	—	—
A55	—	0.0477	0.0532	0.0634	—	—	—	—
A6	—	0.0591	0.0812	0.0533	—	—	—	—
A38	—	0.0521	0.0366	0.0510	—	—	—	—
N4 - N5	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.11: Average discrimination error (ADE) for each pairwise LFC comparison between N4 and N5 calls randomly selected from all available N4 and N5 calls of A pod individuals. Individuals being compared are listed vertically and horizontally. See legend of Table 5.2 for further details.

Forty-two comparisons were possible for the LFC of call types N4 and N9: thirty between MU, six within MU, and six self (Table 5.12). When grouped by comparison type, mean ADE can be ranked in the following order from least to greatest: within MU, between MU and self (Table 5.15).

A8	<i>0.0006</i>	0.0575	0.0343	0.0234	—	0.0096	—	0.0150
A32	0.0522	<i>0.0393</i>	0.0206	0.0061	—	0.0319	—	0.0104
A37	0.0657	0.0401	<i>0.0586</i>	0.0225	—	0.0186	—	0.0172
A46	0.0782	0.0613	0.0107	<i>0.0272</i>	—	0.0137	—	0.0234
A12	—	—	—	—	—	—	—	—
A55	0.0352	0.0284	0.0007	0.0035	—	<i>0.0014</i>	—	0.0043
A6	0.0323	0.0277	0.0008	0.0086	—	0.0021	—	0.0041
A38	0.0550	0.0479	0.0008	0.0044	—	0.0014	—	<i>0.0587</i>
N4 - N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.12: Average discrimination error (ADE) for each pairwise LFC comparison between N4 and N9 calls randomly selected from all available N4 and N9 calls of A pod individuals. Individuals being compared are listed vertically and horizontally. See legend of Table 5.2 for further details.

Eighteen comparisons were possible between call types N5 and N9: nine between MU, six within MU and three self (Table 5.13). When averaged by comparison type, mean ADE can be ranked in the following order: within MU, self and between MU comparisons (Table 5.15).

A8	---	---	---	---	---	---	---	---
A32	0.0556	0.0668	0.0034	0.0230	---	0.1077	---	0.0325
A37	0.0391	0.0329	0.0007	0.0080	---	0.0602	---	0.0057
A46	0.1185	0.0315	0.0108	0.0120	---	0.0944	---	0.0122
A12	---	---	---	---	---	---	---	---
A55	---	---	---	---	---	---	---	---
A6	---	---	---	---	---	---	---	---
A38	---	---	---	---	---	---	---	---
N5 - N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.13: Average discrimination error (ADE) for each pairwise LFC comparison between N5 and N9 calls randomly selected from all available N5 and N9 calls of A pod individuals. Individuals being compared are listed vertically and horizontally. See legend of Table 5.2 for further details.

Within Shared Repertoire – Low Frequency Component – W pod

Six comparisons were performed between call types N32 and N33, the only comparison possible between W pod call types: four within MU and two self comparisons (Table 5.14). When averaged by comparison type, mean ADE can be ranked in the following order from least to greatest: self and within MU comparisons (Table 5.15).

W2	0.2877	0.1815	0.4317
W3	0.1384	0.1244	0.2348
W5	---	---	---
N32-N33	W2	W3	W5

Table 5.14: Average discrimination error (ADE) for each pairwise LFC comparison between N32 and N33 calls randomly selected from all available N32 and N33 calls of W pod individuals. Individuals being compared are listed vertically and horizontally. See legend of Table 5.2 for further details.

Within Shared Repertoire – Low Frequency Component – Summary

When the results of comparisons between different call types of the same repertoire were averaged according to comparison type, between MU comparisons produced the lowest mean (0.0477 ± 0.0452), within MU comparisons had the next highest mean (0.0695 ± 0.0848) and the greatest mean was found for self comparisons (0.0834 ± 0.0806) (Table 5.15, Figure 5.3).

Comparison Type		N2 - N4	N2 - N5	N2 - N9	N4 - N5	N4 - N9	N5 - N9	N32-N33	Total
Btw MU	N	20	6	16	12	29	9	---	92
	Mean (SD)	0.024 (0.023)	0.016 (0.013)	0.113 (0.050)	0.064 (0.022)	0.024 (0.022)	0.058 (0.041)	---	0.0478 (0.045)
	W (<i>p</i>)	0.849 (0.005)	0.927 (0.558)	0.934 (0.287)	0.878 (0.082)	0.902 (0.011)	0.937 (0.549)	---	
Within MU	N	5	4	5	6	7	6	4	37
	Mean (SD)	0.043 (0.037)	0.018 (0.021)	0.105 (0.056)	0.083 (0.040)	0.023 (0.021)	0.018 (0.013)	0.247 (0.130)	0.070 (0.085)
	W (<i>p</i>)	0.869 (0.261)	0.725 (0.022)	0.927 (0.577)	0.826 (0.100)	0.887 (0.257)	0.894 (0.340)	0.880 (0.337)	
Self	N	3	2	3	3	6	3	2	22
	Mean (SD)	0.085 (0.022)	0.002 (0.001)	0.194 (0.018)	0.106 (0.064)	0.031 (0.026)	0.026 (0.035)	0.206 (0.116)	0.083 (0.081)
	W (<i>p</i>)	0.839 (0.213)	---	0.826 (0.177)	0.824 (0.174)	0.869 (0.223)	0.874 (0.307)	---	
	F (<i>p</i>)	0.500 (0.613)	0.567 (0.586)	1.652 (0.216)	1.082 (0.360)	0.368 (0.694)	2.192 (0.146)	0.005 (0.947)	

Table 5.15: Summary of similarity levels measured by neural network performance for between call type LFC comparisons. Values show sample size (N), mean, standard deviation (SD), and results of the Shapiro-Wilks (W, *p*) and Levene's (F, *p*) tests grouped by comparison type (listed vertically) and call type (listed horizontally).

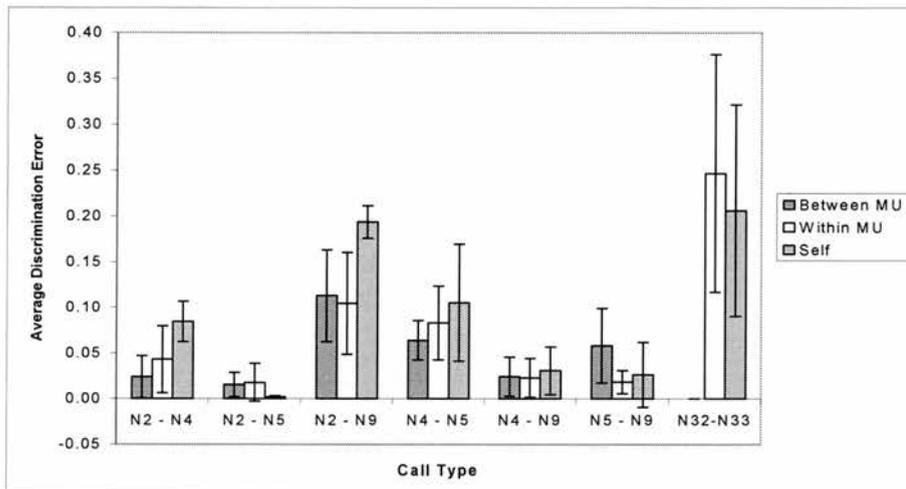


Figure 5.3: Similarity levels for within repertoire LFC comparisons for A and W pod repertoires grouped by comparison type. Error bars represent standard deviation.

High Frequency Component

Calls were selected from the same individuals for all of the following comparisons in order to include the maximum number of individuals for which enough calls were recorded in the analysis. Calls of type N2 were used from A8, A12, A32 and A46 (three matriline) while calls of type N4 were used from A6, A8, A32, A37, A38, A46 and A55 (four matriline). For call type N5 calls were available from the three individuals of matriline A36: A32, A37 and A46. There were N9 calls from A8, A32, A37, A38, A46 and A55, representing four matriline. For W pod comparisons, calls of type N32 were selected from W2 and

W3. All three individuals of the same matriline were included in the analysis of call type N33.

Within Shared Repertoire – High Frequency Component – A pod

Twenty-eight HFC comparisons were possible between call types N2 and N4: twenty between MU, five within MU and three self (Table 5.16). Mean ADE could be ranked by comparison type in the following order from least to greatest: between MU, within MU and self comparisons. Table 5.23 contains means, standard deviation and sample sizes for this and subsequent comparisons.

A8	0.0745	0.0341	0.0388	0.0302	---	0.0064	0.0332	0.0131
A32	0.0444	0.0226	0.0383	0.0226	---	0.0352	0.0225	0.0079
A37	---	---	---	---	---	---	---	---
A46	0.0735	0.0578	0.0571	0.0611	---	0.0046	0.0302	0.0144
A12	0.0614	0.0314	0.0335	0.0387	---	0.0383	0.0363	0.0159
A55	---	---	---	---	---	---	---	---
A6	---	---	---	---	---	---	---	---
A38	---	---	---	---	---	---	---	---
N2 - N4	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.16: Average discrimination error (ADE) for each pairwise HFC comparison between N2 and N4 calls randomly selected from all available N2 and N4 calls of A pod individuals. See legend of Table 5.2 for further details.

Twelve HFC comparisons were possible to compare call types N2 and N5: six between MU, four within MU and two self (Table 5.17). When averaged by comparison type, mean ADE could be ranked in the following order from least to greatest: within MU, between MU and self comparisons (Table 5.23).

A8	----	0.3398	0.2079	0.2840	----	----	----	----
A32	----	<i>0.2691</i>	0.1939	0.1784	----	----	----	----
A37	----	----	----	----	----	----	----	----
A46	----	0.2548	0.2297	<i>0.2691</i>	----	----	----	----
A12	----	0.1682	0.1442	0.1556	----	----	----	----
A55	----	----	----	----	----	----	----	----
A6	----	----	----	----	----	----	----	----
A38	----	----	----	----	----	----	----	----
N2 - N5	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.17: Average discrimination error (ADE) for each pairwise HFC comparison between N2 and N5 calls randomly selected from all available N2 and N5 calls of A pod individuals. See legend of Table 5.2 for further details.

Twenty-four HFC comparisons were possible between call types N2 and N9: sixteen between MU, five within MU comparisons and three self (Table 5.18). Averaged by comparison type, mean ADE could be ranked from least to greatest in the following order: between MU, within MU and self comparisons (Table 5.23).

A8	<i>0.1570</i>	0.1652	0.0721	0.1780	----	0.2514	----	0.0626
A32	0.0750	<i>0.1463</i>	0.0895	0.1530	----	0.1469	----	0.0533
A37	----	----	----	----	----	----	----	----
A46	0.0416	0.0853	0.0346	<i>0.1449</i>	----	0.1679	----	0.0108
A12	0.0334	0.0951	0.0454	0.1570	----	0.1395	----	0.0010
A55	----	----	----	----	----	----	----	----
A6	----	----	----	----	----	----	----	----
A38	----	----	----	----	----	----	----	----
N2 - N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.18: Average discrimination error (ADE) for each pairwise HFC comparison between N2 and N9 calls randomly selected from all available N2 and N9 calls of A pod individuals. See legend of Table 5.2 for further details.

For call-types N4 and N5 there were twenty-one HFC comparisons possible: twelve between MU, six within MU and three self (Table 5.19). Mean ADE could be ranked by comparison type in the following order from least to greatest: between MU, within MU and self comparisons (Table 5.23).

A8	----	0.1324	0.0953	0.0836	----	----	----	----
A32	----	<i>0.1145</i>	0.0759	0.0782	----	----	----	----
A37	----	0.1061	<i>0.0830</i>	0.1002	----	----	----	----
A46	----	0.1033	0.0443	<i>0.1060</i>	----	----	----	----
A12	----	----	----	----	----	----	----	----
A55	----	0.0531	0.0196	0.0117	----	----	----	----
A6	----	0.1016	0.0639	0.0528	----	----	----	----
A38	----	0.0859	0.0427	0.0314	----	----	----	----
N4 - N5	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.19: Average discrimination error (ADE) for each pairwise HFC comparison between N4 and N5 calls randomly selected from all available N4 and N5 calls of A pod individuals. See legend of Table 5.2 for further details.

Forty-two HFC comparisons were possible between call types N4 and N9 calls; twenty-nine between MU, seven within MU and six self (Table 5.20). Comparison types can be ranked in the following order from least to greatest based on mean ADE: between MU, within MU and self comparisons (Table 5.23).

A8	<i>0.0752</i>	0.0328	0.0420	0.0520	----	0.0301	----	0.0168
A32	0.0841	<i>0.0679</i>	0.0286	0.0655	----	0.0654	----	0.0292
A37	0.0510	0.0644	<i>0.0145</i>	0.0515	----	0.0467	----	0.0353
A46	0.1078	0.0895	0.0486	<i>0.0678</i>	----	0.0222	----	0.0314
A12	----	----	----	----	----	----	----	----
A55	0.0615	0.0631	0.0298	0.0319	----	<i>0.0729</i>	----	0.0057
A6	0.0446	0.0488	0.0061	0.0339	----	0.0229	----	0.0183
A38	0.0563	0.0370	0.0062	0.0235	----	0.0223	----	<i>0.0205</i>
N4 - N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.20: Average discrimination error (ADE) for each pairwise HFC comparison between N4 and N9 calls randomly selected from all available N4 and N9 calls of A pod individuals. See legend of Table 5.2 for further details.

Eighteen comparisons were possible between calls of type N5 and N9; nine between MU, six within MU and three self (Table 5.21). Averaging these comparisons by comparison type produced mean ADEs ranked in order from lowest to highest; within MU comparisons, between MU comparisons and self comparisons (Table 5.23).

A8	----	----	----	----	----	----	----	----
A32	0.1352	0.2259	0.1757	0.2307	----	0.3206	----	0.1822
A37	0.0815	0.0837	<i>0.1642</i>	0.1493	----	0.1507	----	0.0943
A46	0.0925	0.1506	0.1170	<i>0.1765</i>	----	0.2743	----	0.1084
A12	----	----	----	----	----	----	----	----
A55	----	----	----	----	----	----	----	----
A6	----	----	----	----	----	----	----	----
A38	----	----	----	----	----	----	----	----
N5 - N9	A8	A32	A37	A46	A12	A55	A6	A38

Table 5.21: Average discrimination error (ADE) for each pairwise HFC comparison between N5 and N9 calls randomly selected from all available N5 and N9 calls of A pod individuals. See legend of Table 5.2 for further details.

Within Shared Repertoire – High Frequency Component – W pod

Due to the small membership of W pod and the limited repertoire available, only six comparisons were possible between call types N32 and N33; four within MU and two self (Table 5.22). Mean ADE could be ranked by comparison type in the following order from least to greatest: within MU and self comparisons (Table 5.23).

W2	<i>0.4108</i>	0.3242	0.2436
W3	0.4218	<i>0.3884</i>	0.1854
W5	----	----	----
N32 - N33	W2	W3	W5

Table 5.22: Average discrimination error (ADE) for each pairwise HFC comparison between N32 and N33 calls randomly selected from all available N32 and N33 calls of W pod individuals. See legend of Table 5.2 for further details.

Within Shared Repertoire – High Frequency Component – Summary

For HFC comparisons between two call types of a shared repertoire, there were one hundred and fifty-one comparisons: ninety-two between MU, thirty-seven within MU and twenty-two self comparisons. Unlike comparisons within a single call type, ADE is affected more by the call types compared than by comparison type. Between-MU comparisons still had the lowest mean ($0.0741 \pm$

0.0720), within MU comparisons had the next lowest mean (0.1224 ± 0.0908) and self comparisons had the highest mean (0.1424 ± 0.1105) (Table 5.23, Figure 5.4).

Comparison Type		N2 - N4	N2 - N5	N2 - N9	N4 - N5	N4 - N9	N5 - N9	N32-N33	All CTs
Btw MU	N	20	6	16	12	29	9	---	92
	Mean (SD) W (p)	0.030 (0.018) 0.929 (0.149)	0.217 (0.079) 0.879 (0.263)	0.097 (0.072) 0.929 (0.232)	0.064 (0.013) 0.973 (0.939)	0.039 (0.023) 0.934 (0.068)	0.160 (0.085) 0.847 (0.069)	---	0.074 (0.072)
Within MU	N	5	4	5	6	7	6	4	37
	Mean (SD) W (p)	0.043 (0.015) 0.889 (0.351)	0.214 (0.035) 0.954 (0.742)	0.100 (0.047) 0.938 (0.655)	0.085 (0.024) 0.872 (0.234)	0.052 (0.024) 0.971 (0.908)	0.151 (0.050) 0.976 (0.929)	0.294 (0.103) 0.980 (0.899)	0.122 (0.091)
Self	N	3	2	3	3	6	3	2	22
	Mean (SD) W (p) F (p)	0.053 (0.027) 0.927 (0.479) 0.263 (0.771)	0.269 (0) --- 2.261 (0.129)	0.149 (0.007) 0.839 (0.212) 1.652 (0.216)	0.101 (0.016) 0.934 (0.502) 1.680 (0.214)	0.053 (0.028) 0.747 (0.018) 0.043 (0.958)	0.189 (0.033) 0.892 (0.361) 0.992 (0.394)	0.400 (0.016)	0.142 (0.110) 3.614 (0.130)

Table 5.23: Summary of network performance for between call type comparisons of the high frequency component within a repertoire. Values show sample size (N), mean, standard deviation (SD) and results of Shapiro-Wilks (W, p) and Levene's (F, p) tests for comparisons grouped by comparison type (listed vertically) and call type (listed horizontally).

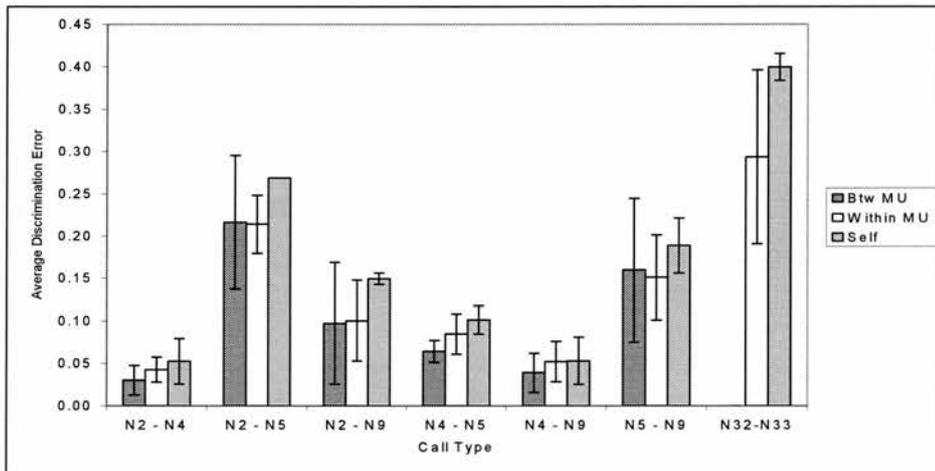


Figure 5.4: Average discrimination error for HFC comparisons at the within repertoire level. Comparisons were performed between two different call types within the repertoires of A pod and W pod. Error bars represent standard deviation.

Within Shared Repertoire – Summary

Call type and comparison type each had a statistically significant effect on similarity level, although this effect was not altered when the Bonferroni correction was applied (call type: $F_{6,262}=24.719$, $p<0.05$; comparison type: $F_{2,262}=7.46$, $p<0.05$). The Bonferroni correction did result in significant similarity levels between components ($F_{1,262}=67.94$, $p<0.05$). There was no interaction between

call type and comparison type ($F_{11,262}=1.23, p=0.26$), or comparison type and component ($F_{1,262}=1.92, p=0.17$), but there was a significant interaction between component and call type ($F_{11,262}=28.522, p<0.05$). There was no interaction between all three factors ($F_{12,262}=1.23, p=0.26$).

Unshared Repertoire Analysis

Low Frequency Component

One hundred comparisons were performed using the LFC of single A pod and W pod call types to evaluate the similarity between low-frequency components of calls from different repertoires. Because these comparisons were between individuals with a more infrequent level of social association and genetic relatedness, they were described as the ‘between clan’ comparison type.

Unshared Repertoire – Low Frequency Component – A and W pods

The number of high quality calls for each call type limited the total number of comparisons possible. For all comparisons including N32 calls, these calls were selected from individuals W2 and W3 from the same matriline. N33 calls were also only available from a single matriline and were from W2, W3 and W5. Comparisons using call type N2 included calls from individuals A8, A12, A32 and A46, representing four matriline. Calls of type N4 were available from four matriline and included individuals A6, A8, A32, A37, A38, A46 and A55. Only one matriline had N5 calls for comparison and these were from individuals A32, A37 and A46. Finally, enough calls of type N9 were recorded from A8, A32, A37, A38, A46 and A55, including calls from four matriline. Similarity

matrices as well as mean ADE for all between clan comparisons are shown in

Table 5.24 and Figure 5.5.

Call Type		Average Discrimination Error											
N32 - N2	W2	0.0517	0.0185	---	0.0118	0.0304	---	---	---	8	N		
	W3	0.0643	0.0407	---	0.0492	0.0447	---	---	---			0.039 (0.018)	Mean (SD)
	W5	---	---	---	---	---	---	---	---				
N32 - N4	W2	0.0128	0.0099	0.0027	0.0118	---	0.0023	0.0049	0.0037	14	N		
	W3	0.0597	0.0483	0.0409	0.0532	---	0.0547	0.0449	0.0458			0.028 (0.023)	Mean (SD)
	W5	---	---	---	---	---	---	---	---				
N32 - N5	W2	---	0.0101	0.0113	0.0095	---	---	---	---	6	N		
	W3	---	0.0515	0.0543	0.0527	---	---	---	---			0.032 (0.0233)	Mean (SD)
	W5	---	---	---	---	---	---	---	---				
N32 - N9	W2	0.0297	0.0355	0.0209	0.0158	---	0.0112	---	0.0165	12	N		
	W3	0.0465	0.0582	0.0513	0.0461	---	0.0444	---	0.0598			0.36 (0.017)	Mean (SD)
	W5	---	---	---	---	---	---	---	---				
N33 - N2	W2	0.0104	0.0293	---	0.0046	0.0399	---	---	---	12	N		
	W3	0.0788	0.0483	---	0.0504	0.0441	---	---	---			0.049 (0.025)	Mean (SD)
	W5	0.0734	0.0567	---	0.0733	0.0755	---	---	---				
N33 - N4	W2	0.0322	0.0160	0.0372	0.0050	---	0.0028	0.0196	0.0181	21	N		
	W3	0.0365	0.0405	0.0321	0.0309	---	0.0220	0.0189	0.0219			0.034 (0.018)	Mean (SD)
	W5	0.0642	0.0611	0.0581	0.0562	---	0.0430	0.0396	0.0488				
N33 - N5	W2	---	0.0218	0.0219	0.0109	---	---	---	---	9	N		
	W3	---	0.0285	0.0211	0.0339	---	---	---	---			0.035 (0.020)	Mean (SD)
	W5	---	0.0671	0.0554	0.0583	---	---	---	---				
N33 - N9	W2	0.0728	0.0326	0.0449	0.0023	---	0.0487	---	0.0428	18	N		
	W3	0.0852	0.0862	0.0376	0.0489	---	0.0514	---	0.0430			0.063 (0.028)	Mean (SD)
	W5	0.1106	0.1067	0.0856	0.0679	---	0.0700	---	0.0724				
		A8	A32	A37	A46	A12	A55	A6	A38	1.418 (0.207)	F (p)		

Table 5.24: Summary of average discrimination error for all between call type LFC comparisons made between individuals from different clans. Calls of W pod individuals (listed vertically) were compared with calls from A pod individuals (listed horizontally). Values in the right-hand column show sample size (N), mean, standard deviation (SD), and results of the Shapiro-Wilks test.

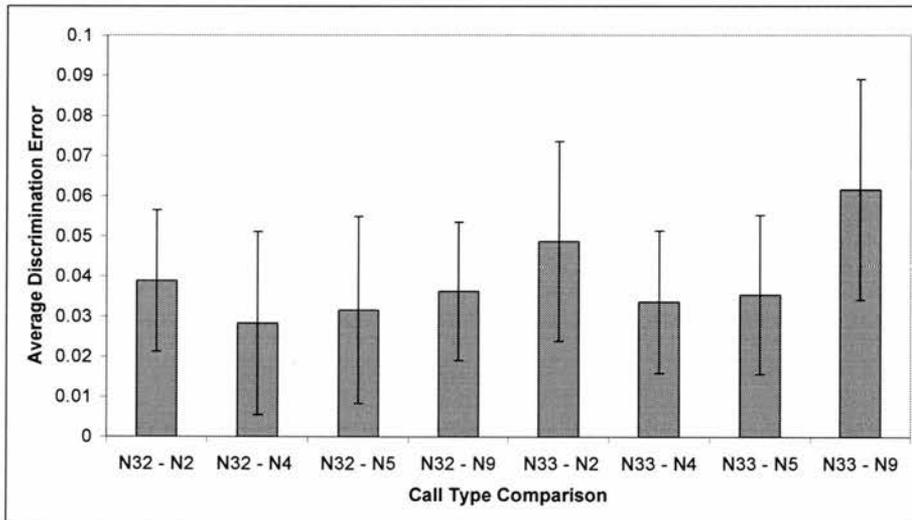


Figure 5.5: Similarity levels of LFC comparisons between call types from different repertoires. Bars represent average discrimination error for all comparisons between single call types from A pod and W pod. Error bars represent standard deviation.

High Frequency Component

Unshared Repertoire – High-Frequency Component – A and W pods

Comparisons of two call types originating from different acoustic clans produced one hundred comparisons when analyzing similarity of the high-frequency component. Because these comparisons do not fit into one of the three existing comparison types (between MU, within MU or self comparisons), they will be described as belonging to comparison type ‘between clan’. All of the following comparisons include individuals from two separate clans and were performed on HFC contours.

Comparisons including call type N2 used calls from individuals A8, A12, A32 and A46 that were from three different matriline. Enough N4 calls were available to perform comparisons across four matriline and involved the following individuals: A6, A8, A32, A37, A38, A46 and A55. N5 calls were only recorded from a single matriline, so comparisons were only with A32, A37 and A46. The following individuals from four different matriline produced enough N9 calls to be compared: A8, A32, A37, A38, A46 and A55. Comparisons of type N32 calls were possible only for within a single matriline (individuals W2 and W3). N33 calls used were from individuals W2, W3 and W5 from the same matriline. Table 5.25 and Figure 5.6 contain the mean result and number of comparisons possible for each between call type comparison.

Call Types		Average Discrimination Error											
N32 - N2	W2	0.1046	0.0544	---	0.0439	0.0389	---	---	---	8	N		
	W3	0.0284	0.0033	---	0.0049	0.0017	---	---	0.035 (0.035)			Mean (SD)	
	W5	---	---	---	---	---	---	---	(0.880) 0.187			W (p)	
N32 - N4	W2	0.0469	0.0169	0.0478	0.0256	---	0.0100	0.0298	0.0146	14	N		
	W3	0.0241	0.0097	0.0405	0.0059	---	0.0090	0.0100	0.0034			0.021 (0.015)	Mean (SD)
	W5	---	---	---	---	---	---	---	---			0.881 (0.059)	W (p)
N32 - N5	W2	---	0.0917	0.0919	0.0950	---	---	---	---	6	N		
	W3	---	0.0404	0.0479	0.0244	---	---	---	---			0.065 (0.031)	Mean (SD)
	W5	---	---	---	---	---	---	---	---			0.832 (0.111)	W (p)
N32 - N9	W2	0.0402	0.0692	0.0568	0.0481	---	0.0784	---	0.0600	12	N		
	W3	0.0357	0.0350	0.0019	0.0032	---	0.0115	---	0.0160			0.038 (0.026)	Mean (SD)
	W5	---	---	---	---	---	---	---	---			0.949 (0.625)	W (p)
N33 - N2	W2	0.0374	0.0133	---	0.0348	0.0082	---	---	---	12	N		
	W3	0.1119	0.1098	---	0.1115	0.0863	---	---	---			0.070 (0.038)	Mean (SD)
	W5	0.0923	0.0939	---	0.0787	0.0626	---	---	---			0.892 (0.126)	W (p)
N33 - N4	W2	0.0301	0.0078	0.0312	0.0193	---	0.0022	0.0097	0.0029	21	N		
	W3	0.0241	0.0078	0.0557	0.0284	---	0.0215	0.0222	0.0156			0.018 (0.014)	Mean (SD)
	W5	0.0383	0.0247	0.0207	0.0062	---	0.0010	0.0125	0.0034			0.927 (0.118)	W (p)
N33 - N5	W2	---	0.0628	0.0782	0.0491	---	---	---	---	9	N		
	W3	---	0.1478	0.1685	0.1455	---	---	---	---			0.096 (0.045)	Mean (SD)
	W5	---	0.0735	0.0658	0.0731	---	---	---	---			0.816 (0.031)	W (p)
N33 - N9	W2	0.0329	0.0305	0.0233	0.0340	---	0.0354	---	0.0373	18	N		
	W3	0.0633	0.0759	0.0372	0.0815	---	0.0883	---	0.0652			0.046 (0.021)	Mean (SD)
	W5	0.0400	0.0416	0.0169	0.0589	---	0.0306	---	0.0387			0.896 (0.049)	W (p)
		A8	A32	A37	A46	A12	A55	A6	A38			0.943 (0.120)	F (p)

Table 5.25: Summary of similarity measured through neural network performance for all between call-type HFC comparisons made with call types from two different clans. Calls from W pod individuals (listed vertically) are compared with those of A pod individuals (listed horizontally). See legend of Table 5.2 for further details.

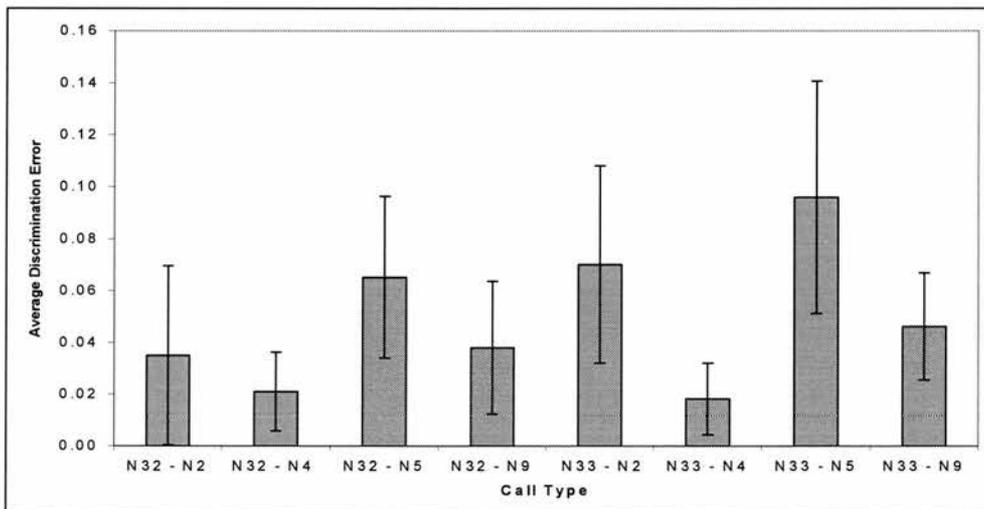


Figure 5.6: Similarity levels for HFC comparisons between call types from different clans. Bars represent average discrimination error for all comparisons between one call type from a member of A pod and one call type from a member of W pod. Error bars represent standard deviation. See Table 5.25 for mean values.

Unshared Repertoire – Summary

All comparisons performed between call types from two different repertoires were of the same comparison type (between clan comparisons), so statistical tests were only used to determine the influence of call type on the level

of similarity. Discrimination error was statistically significantly different between call type ($F_{7,184}=10.30, p<0.05$) and component ($F_{1,184}=6.38, p<0.05$).

Additionally, there was a significant interaction between the factors ($F_{7,184}=6.49, p<0.005$). Due to this interaction, post-hoc tests could not be performed.

Comparison of Repertoire Analyses

Each call type was involved in three sets of comparisons, each on a different repertoire level. First, calls were compared with other calls of the same type. The second set of comparisons was between different call types from a repertoire of calls share within a clan repertoire and finally, call types from different clan repertoires were compared. The results of these comparisons was included in the previous three sections of this chapter, but has been summarized in the figures below for the low frequency component (Figure 5.7) and the high frequency component (Figure 5.8).

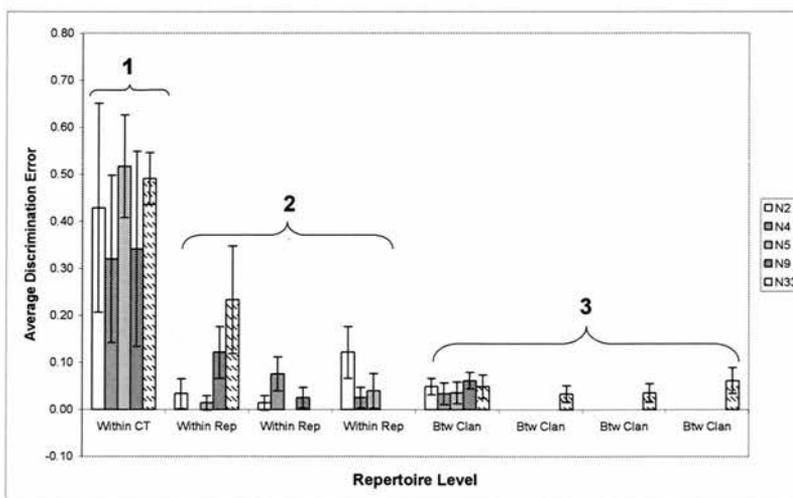


Figure 5.7: Similarity levels of all LFC comparisons from the following three repertoire levels; 1) same call types of the same repertoire, 2) different call types of the same repertoire and 3) different call types of different repertoires. Brackets delineate comparisons from the same repertoire level and separate groups within repertoire level show comparisons of each call type.

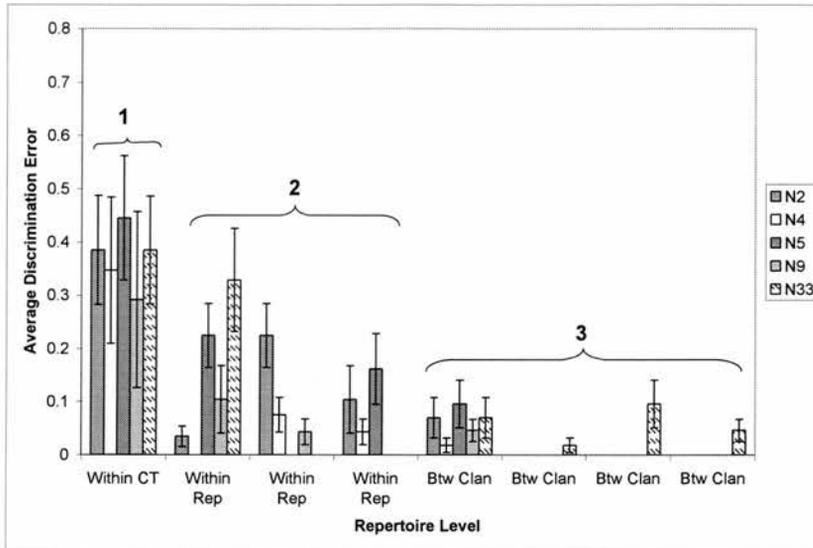


Figure 5.8: Similarity levels of all HFC comparisons from the following three repertoire levels; 1) same call types of the same repertoire, 2) different call types of the same repertoire and 3) different call types of different repertoires. Brackets delineate comparisons from the same repertoire level and separate groups within repertoire level show comparisons of each call type.

For this final analysis, the influence of call type, component and repertoire level on call similarity was compared. Comparison types differed according to repertoire level so they were not evaluated across the three analyses. Significant interactions occurred between all three factors ($F_{16,589}=3.98, p<0.05$) and between call type and repertoire level ($F_{11,589}=15.12, p<0.05$). Similarity levels were significantly different between repertoire levels ($F_{2,589}=6.38, p<0.05$) but not between call types ($F_{6,589}=1.84, p=0.09$). Component type did not interact with either call type ($F_{2,589}=0.91, p=0.09$) or repertoire level ($F_{1,589}=0.17, p=0.68$), and did not significantly affect similarity ($F_{1,589}=0.17, p=0.68$). Bonferroni corrections caused no changes in significant factors.

DISCUSSION

Patterns within call type comparisons

Both the low and high frequency components demonstrate a clear trend that shows the correlation between comparison type and vocal similarity. Calls that were selected from individuals of different matrilineal units are clearly distinguishable, while calls that are from individuals of the same matrilineal unit are much more similar. Calls from the same individual are highly similar, and no easier to distinguish than by chance as the proportion classified correctly and incorrectly was approximately 50%. The grouping of these results by social affiliation demonstrates that social structure does have a strong influence on vocal behavior.

These results also showed that the amount of group-specific signature varies according to call type; there was little group-specific information in N2 calls, but more for N4 and N9 calls. Although this trend is most strongly evident in the HFC results, suggesting that the HFC of N2 calls is less group-specific than the HFC of the other two call types. Additionally, the lack of a strong correlation between LFC and HFC results for the within MU comparison type suggests that potentially individually identifiable information is more strongly conveyed by one component than the other.

Patterns within shared repertoire comparisons

Similarity between call types was much lower than that of calls of the same type, validating the procedure of classifying calls to type by visual

inspection (Ford and Fisher 1983; Ford 1987; Ford 1991) (Figures 5.3 and 5.4). This trend was clearly visible in both the LFC and HFC results. The small standard deviation within comparisons of the same two call types showed that the level of social affiliation between the individuals that produced them does not influence their level of similarity. Comparisons of the same call type had a high standard deviation due to the influence of individual interactions; the more constant the social association, the more similar their calls will be. Additionally, the significant difference between components suggests that both components play a role in distinguishing between call types.

However, some call types did appear to have more similar contours than others and this could be an example of call types that have diverged more recently than other, or call types that are used in the same behavioral interactions. For example, the high-frequency components of call types N2 and N5 were more similar than any of the other comparisons between two call types, and reached the level of similarity found for within call type comparisons. As these are the two call types most frequently recorded during inter-pod associations (Ford 1989), they could transfer a signal regardless of the type of affiliation between the calling individuals. Another interesting result from the series of within repertoire comparisons was the high level of similarity between the two call types of W pod. Call types N32 and N33 had a level of similarity that is close to that of the within call type comparisons. Although these were the only two call types available from W animals, their repertoire does include several more call types (Ford, 1987). However, W pod belongs to a very small clan, R clan, which consists of

two pods, six matriline and 32 individuals in total. In comparison, A clan consists of 10 pods, 19 matriline, and 108 animals (Ford et al, 2000). The fact that the repertoire is shared by a smaller number of individuals could reduce the need for very diverse call types. Calls from more individuals of this clan are needed before any conclusions can be made about their vocal repertoire.

Patterns within unshared repertoire comparisons

The results of comparisons between individuals from different clans showed that calls from different repertoires were more consistently different than calls from the same repertoire, in line with the theory that individuals in different clans are the descendants of one ancestral pod and therefore have independently developed repertoires that have very little acoustic similarity (Ford 1991; Ford *et al.* 2000). Further comparisons between several different repertoires are necessary to confirm such ancestral splits. The results of comparing calls from identified individuals confirmed that previous studies which used calls recorded from groups of individuals (Ford 1989; Deecke *et al.* 1999; Miller and Bain 2000) presented a valid interpretation of individual interactions as grouped by social association. None of these call type combinations are used consistently during a behavioral interaction, so all of the pairs of call types differ to roughly the same extent. These comparisons provide another level to demonstrate the strong influence of social structure on vocal similarity. There are many features that allow for the discrimination of vocalizations between these completely genetically separated individuals.

Patterns across repertoires

The significant interaction between call types, components and repertoire types indicates the complexity of acoustic communication in killer whales. Although the high level of interaction between these three variables limits the extent to which trends are visible in these results, it does support the observations of previous work in the same area. The interaction between repertoire and call type supports the practice of dividing call types into repertoires, as the acoustic distance present between repertoires is related to the lack of similarity between call types (Ford 1987; Ford 1991). Repertoire level does influence call similarity, as has been documented by the documented correlation between acoustic similarity and social affiliation at the clan, pod and matrilineal unit level (Ford 1991; Deecke 2000; Miller and Bain 2000; Yurk *et al.* 2002). Finally the lack of significance between components should be noted. Although similarity levels were measured separately for each component, there was no clear difference between them, suggesting that they function as a discrete unit. A previous study of killer whale vocalizations found that both components independently contained features distinctive to matrilineal units (Miller and Bain 2000).

Signature system of resident killer whales

These comparisons demonstrated that calls are distinguishable between individuals as well as the level of the group, reflecting the importance of both individual and group bonds within killer whales. LFC comparisons showed that the similarity of calls from individuals of different groups was less than that of

individuals from the same group. While the network discriminated calls from the same individual no better than by chance, calls of different family members within the same matrilineal unit were distinguishable. Individual signatures may be important for animals living in the marine environment as they mitigate many of the difficulties of a low-visibility, feature-less habitat that is constantly changing. One well documented example is found within groups of bottlenose dolphins (*Tursiops truncatus*). Individuals have specific 'signature whistles' which identify them to other individuals (Tyack 1997; Janik and Slater 1998).

Variation in similarity levels between call types suggests that these calls constitute a more complex signature system than previously reported. These differences could suggest a benefit to transmitting information in different ways and demonstrate functions for different call types. Killer whales group together for social activities, but then disperse while foraging (Ford 1989; Hoelzel 1993). Each of these behaviors places specific demands on the communication necessary for interactions between individuals as well as interactions between groups as a whole. Individually distinctive calls may be useful in behavioral situations where interactions between individuals are important. Inter-pod associations frequently occur and include multiple matrilineal units located at close range (Heimlich-Boran 1986; Osborne 1986; Yurk *et al.* 2002). If group-specific calls act as a password to coordinate movements only within group members (Feeles 1982), usage of these calls would be less appropriate for inter-pod meetings since there is no reason to exclude or include certain individuals. Although the high-frequency component attenuates more quickly than the low-frequency component, it is much

easier to localize. For animals foraging at distance and attempting to maintain group cohesion, receiving a highly directional signal would allow them to calculate the orientation and heading of the calling individual (Miller 2002). Vocal response to that caller would then confirm receipt of that directional information as well as indicate the receiver's location. This could explain the function of matched call-type vocal exchanges in this species (Miller *et al.* 2004b).

Additionally, the nested effect of combining highly recognizable group signatures with less prominent individual signatures allows for both more possible signatures as well as more efficient discrimination. It has been suggested that some bats with group-specific calls use a combination of signature systems (scent cues and/or relative location combined with acoustic identification) to identify individuals at close range (McCracken 1993; Boughman 1997). In the marine environment, the senses of both smell and sight are very limited, making identification by sound the most likely way to confirm an individual's identity. In this way, group-specific high frequency calls may be used to weed out important signals (those from a group member) from the background calls of non-group members. Once a listening individual has determined that the caller is from the same group and therefore transmitting information relevant to the listener, it can localize the position of that calling individual, extract the individual's specific identity and respond accordingly.

The relatively small size of killer whale groups should also make it easier for individuals to be correctly identified. The probability of correctly classifying

a target individual using signature information is directly related to the number of individuals that the target individual could possibly be (Beecher 1989). Maintaining different levels of signature systems would reduce the potential number of individuals it could be. For example, instead of discriminating between all 216 northern resident whales, group-specific signatures would allow an individual of A pod to discriminate between the sixteen group-specific signatures of A clan. Individually specific signatures would then be needed to discriminate between the 3-6 possible individuals within that matriline, greatly increasing the likelihood of correct identification through reducing the number of possible individuals.

The limited distribution of these animals along with their long lifespan also reduces the chances they will encounter an unfamiliar group signature. This factor makes the signature system used by killer whales very different from that of colonially living bats and birds, both of which use individually-specific vocalizations to identify individuals (Gelfand and McCracken 1986; Beecher 1989; Balcombe and McCracken 1992; Medvin *et al.* 1992; McCracken 1993).

Social learning

This study also demonstrated new evidence in support of the theory that vocalizations are learned from long-term associations between individuals. As these results showed, the high-frequency component of call type N2 is much more similar between different matrilineal units than either call type N4 or N9. This could be explained by the observation that call type N2 is most frequently used in

inter-pod associations where many animals are present (Ford 1989). During these meetings, an individual will hear animals from different matriline using this call type, so matriline are less likely to develop it independently. On the other hand, matrilineal units traveling on their own generally use N9 calls (Ford 1989). Individuals of a matrilineal unit will only hear the N9 version of their own matriline, so any variations added by an individual will be propagated within that group of individuals (Ford 1989), creating highly group-specific calls. The higher similarity level between call types N2 and N5 may also be connected to the observation that these two call types are used more frequently together in situations where multiple matrilineal units are present.

Social learning, or the ability for individuals to acquire skills and behaviors from the individuals with whom they spend the most time, has been documented in several other species of cetacean and has been suggested as a mechanism for sharing of information in killer whales (Boran and Heimlich 1999; Deecke 2000; Miller and Bain 2000; Yurk *et al.* 2002). As these results show, individuals that constantly travel together have vocalizations that are much more similar than those of animals that are not socially affiliated. The different feeding specializations of resident and transient whales, as well as the techniques of catching sea lions by stranding (Lopez and Lopez 1985; Guinet 1991; Guinet and Bouvier 1995) and 'carousel' feeding around schools of herring (Similä and Ugarte 1993; Domenici *et al.* 2000) are further evidence for social learning within this species (Tyack and Sayigh 1997; Boran and Heimlich 1999). Highly similar vocalizations between affiliated but unrelated individuals also can demonstrate

learning between individuals. However, as this study looks at the similarity of calls between highly related individuals, it does not confirm the presence of social learning any further than is already conjectured.

Important considerations

Although this study documented consistent features in the frequency contour of calls from individuals, there are still several important aspects that should be considered, especially before further studies of this topic are made. First of all, these features were identified by computer analysis and not by the animals themselves. Before evidence for individual signatures can be definitively determined, the ability of killer whales to perceive individual signature information should be tested using playbacks of calls to individuals, preferably those in the wild. Behavioral responses of wild cetaceans are very difficult to measure, but the recent development of a detachable tag that records both animal orientation and vocalizations would make it possible to record any reactions of the animal to playbacks (Johnson and Tyack 2003). This tag has already been demonstrated to successfully (and non-invasively) record the movement of cetaceans in their natural environment with killer whales and several other species of cetacean (Miller *et al.* 2004a; Nowacek *et al.* 2004). Modifications of calls used in the playback, such as using calls with only one frequency component or altered frequency modulations, may also provide information about which features are important in conveying information between animals.

Secondly, this study only measured the level of similarity found in the frequency contour of calls from individuals. These contours have been shown to be features perceived by cetaceans (Ralston and Herman 1995), making it a useful acoustic feature to explore initially for individual signature information. Many other acoustic features, such as amplitude modulations or spectral features could be as important in indicating individual identity as the frequency contour. Before the signature system for identifying individual killer whales can be confirmed, there are many more acoustic features that should be explored. This study demonstrated that the frequency structure of vocalizations possibly contains individually consistent features, but it is important to remember that this could be only a portion of a more complex overall system of identification.

Finally, although these trends were clearly visible for the individuals involved, these individuals represent only a small portion of the resident killer whale population inhabiting the Pacific Northeast. Comparisons were possible across all comparison types for only three call types, which is a very small portion of the entire repertoire used by these animals. Additionally, these call types were 'active' call types, that is call types that are generally used by individuals across a wide variety of high energy behavioral situations such as foraging and traveling (Ford, 1989). None of the 'resting' call types, such as N3, N7 or N8 were compared and trends in vocal similarity between individuals could be different for those call types. Also, the majority of calls used in this study were from the individuals of matriline A36. These individuals, A32, A37 and A46, are all adult males with no other individuals in their matrilineal unit. Trends indicated could

be biased by the relatively atypical social affiliation of this group. Also, two-thirds of all vocalizations were produced by males, which could influence the level of vocal similarity that was measured. Finally, there were no considerations made for ambient noise levels or behavioral states. Comparisons included the largest number of calls available from the maximum number of follows to reduce this element, but a much larger and more evenly distributed sample set is necessary to definitively exclude the influence of such factors and reduce any bias caused by specific individuals. A more comprehensive study of the vocalizations of individuals from more pods and different clans is necessary before these results can be definitively proven to be representative of all resident killer whales. Recording free-ranging animals eliminates the ability to select specific individuals to record and significantly reduces the number of vocalizations from individuals useable for analysis. In light of the results of this study, more time in the field is necessary to make these results conclusive, despite the difficulties inherent in attempting to collect specific vocalizations from specific individual killer whales.

Conclusion

From the point when they are born into a matriline to the end of their lives, killer whales will hear sounds a limited number of individuals, and the identity of those individuals does not change rapidly due to the longevity of this species. Identifying the group to which an individual belongs may be more important for these animals than identifying the exact individuals within that group. However, individually specific features are still present and discrimination between individuals could be possible through a vocal signature, meaning that individual

bonds within matriline are maintained. Further work is necessary to explore individual roles within these matriline and to understand the extent to which these individual bonds influence killer whale interactions and behavior. However, group bonds appear to be much more significant and of such benefit to individuals that individual signature information is greatly overshadowed by the effects of long-term social affiliations.

CHAPTER 6: ***General Conclusion***

Study Goals

The goals of this study were to explore the effect of social relationship on the ability of a neural network to discriminate which individual produced a call, and to specifically investigate the possibility that individually distinctive information is communicated through the time-frequency contours of killer whale calls. Performance of a neural network program for pattern recognition was used to assess the similarity of vocalizations between individual killer whales that associate to various extents. Killer whales have been observed to travel with the same group of individuals, and such group-living appears to be highly beneficial to the individuals involved. As vocalizations are used to coordinate the interactions of individuals within that group, it is very likely that group-living has considerably influenced the vocal behavior of individuals. Comparisons of vocalizations were performed between individuals from different matriline, between individuals in the same matriline and between different calls from the same individual. In order to confirm that results were descriptive of an individual's typical vocal behavior and not an effect of the way calls were selected from the entire dataset, the initial experiment consisted of testing three different methods of call selection before presentation to the neural network.

Chapter 3 detailed the results of the first of these experiments and explored the difference between similarity levels from comparing of two sets of calls selected specifically according to context (described as the 'hand selection' method) and those from averaging the results from many comparisons of calls

selected entirely at random (described as the 'random selection method). Due to the small number of available recording sessions for each individual and the low number of calls produced by most of the individuals, selecting calls specifically by context greatly reduced the available comparisons and only used a small proportion of the available data. Randomly selecting a smaller number of calls many times allowed the inclusion of more individuals within the analysis and used more calls from each individual, creating a sample that was more representative of an individual's range of calling. Network performance did not differ significantly between these two methods. Because the random selection method produced an adequate representation of an individual's calls while requiring a smaller number of calls, it was decided that this method would be used in the final analysis of similarity.

In Chapter 4, a second test of selection method was performed to determine whether there was any influence of group calling bouts on individual calls. Killer whales generally vocalize in matched call-type bouts (Miller *et al.*, 2004), calls within which may not be independent, but are in effect the same call repeated multiple times. Several calls in sequence might have a more similar structure than those produced in an isolated context and therefore would create levels of similarity that describe the calling situation and not the full variability of the call itself. Even if this is true and bouts do influence individual calls, selection by bout did not produce a different level of similarity than random selection from the entire set of available calls. Further work specifically designed to look at individuals calls produced during vocal exchanges would be necessary

to determine if within-bout calls have a more similar structure than those produced in an isolated context.

The final experiment in this study (Chapter 5) determined the level of acoustic similarity present between individuals of different levels of social affiliation and fully explored the extent of similarity within the repertoires of resident killer whales. The distinctive shifts in the pulse-repetition rate of killer whale calls allow them to be classified into discrete call types. Several studies have catalogued and measured the differences between these call types (Ford 1987). For this study, comparisons were first performed between two calls within the same call type for each of the two simultaneous frequency contours they contain (low frequency component and high frequency component) separately. Next, between call type comparisons were made between calls of a single clan and then finally between calls from two different clans. The last two sets of comparisons evaluated similarity at the clan level, which represents the lowest amount of maternal relatedness within resident killer whale social structure.

Signature systems

Group specific

The presence of signature information at different levels in killer whale social structure creates an interesting comparison with the signature systems used by other cetaceans as well as those of other taxa. Throughout the animal world, the use of individually identifiable vocalizations is widespread relative to the number of species that use group-specific vocalizations (Boughman, 1997). The

function of group-specific vocalizations appears to be similar within these species; by creating an easily distinguishable ‘badge’ of membership, individuals can either be included or excluded depending on whether they share the badge. Feekes (1982) termed this function the ‘password hypothesis’ implying that only those individuals that knew the correct calls would be accepted in the group. For group-living species, such as budgerigars (Bartlett & Slater, 1999), and Australian magpies (Farabaugh *et al.*, 1988), this membership may help to defend a territory by immediately identifying and excluding non-group members. More opportunistic foragers like greater spear-nosed bats use group-specific calls to coordinate movements between individuals within the group (Boughman & Wilkinson, 1998). The one life history parameter that is clearly shared by all of these species is their tendency to live in stable groups that have a fixed membership regardless of genetic ties.

Many authors have used vocalizations to measure social interactions between members of a species (Wilkinson & Boughman, 1999; Tyack, 1986). Long-term photo-ID studies have demonstrated that individual killer whales travel, forage and socialize throughout their lives with the same individuals, and that life-long social bonds occur between a mother and her offspring (Ford *et al.*, 2000; Bigg, 1982), an observation which has subsequently been supported by genetic tests (Barrett-Lennard, 2000). The strong social bonds between kin would predict that group-specific vocalizations would predominate in the vocal behavior of killer whales. Several studies using different methods of analysis have indeed confirmed the presence of group-specific vocalizations at different levels of social

association for all of the resident killer whale communities (Ford, 1991; Ford, 1989; Ford, 1987; Miller & Bain, 2000; Deecke *et al.*, 1999; Yurk *et al.*, 2002). These studies have proposed that the function of group-specific signatures in killer whale calls is to facilitate group cohesion while foraging at distance. The finding that these calls also have the potential to act as individual signatures demonstrates that killer whales use a more complex vocal system than previously suggested.

Individual-specific

In species whose foraging strategy requires frequent separations between individuals that are highly dependent on each other for survival, individually distinctive contact calls are essential in reuniting individuals. Individually recognizable calls are especially common between mothers and offspring in species in environments where location, odor and other cues are not possible. Several studies have shown the presence of individuality in the vocalizations of two penguin species (king penguins, *Aptenodytes patagonicus*, and emperor penguins, *Aptenodytes forsteri*), species which lack a fixed nesting site and alternate parents in their care of the young (Lengagne *et al.*, 2001; Robisson *et al.*, 1993). Species of nesting penguins (genus *Eudyptes* and *Pygoscelis*) do not have strong individual vocal signatures (Searby *et al.*, 2004) as they use nest location to reunite within the many individuals of their colony. Similarly, different breeding strategies in pinnipeds also correlate with the presence of individual vocal signatures. Female otariids periodically separate from their pups to forage at sea

before the pup is completely weaned and must be able to identify their offspring on return to the large colony. By contrast, phociid mothers live off stored fat during the breeding season and do not separate from their offspring at any point while on land. Individually recognizable vocalizations have been shown in several species of otariids, but are not known within phocid females (Charrier *et al.*, 2002; Phillips & Stirling, 2000; Trillmich, 1981). Bottlenose dolphin mother/calf pairs also can identify each other with individually specific 'signature whistles' (Sayigh *et al.*, 1998). In a low visibility environment without fixed meeting sites, these vocalizations allow calves to reunite with the only individual that will keep them alive.

Individual killer whales spend periods of time dispersed in order to forage and then reunite without the use of sight, smell or a fixed reunion site. However, for resident killer whales hunting abundant salmon, traveling in a group of several individuals increases foraging success. It has also been suggested that killer whales benefit from group-living by sharing care for young, and that the presence of female group members past their reproductive years serves to assist young individuals and/or pass on knowledge of choice feeding areas (Heimlich-Boran, 1986). All of these reasons show the importance of strong within group social bonds throughout their lifetime yet individual bonds may also be important within these groups. Strong individual associations would also present the need for individually specific identification (between mothers and calves for example), and previous studies have suggested the possibility for individually specific vocalizations in southern resident killer whales (Hoelzel & Osborne, 1986). The

two component structure of stereotyped calls allows both sets of information to be conveyed at the same time.

Future work

Spectrographic analysis

The constraints of this analysis leave several factors still to be checked before features of individuality can be confirmed. First of all, the neural network is useful in analyzing patterns as a single unit, maintaining the signal in the exact form that would be received by an individual in the wild. However, while the network does match similar features between calls, this information cannot be independently extracted from within the network. A different method of analysis that measures individual features, such as a discriminant function analysis, would be needed to identify the specific features that encode individuality.

Nevertheless, the network analysis was useful in analyzing frequency component contours as a single unit. Factors to distinguish between the calls of individuals were measured in each of the two frequency components independently, an unlikely circumstance to arise during killer whale communication in the wild. Further analysis of killer whale vocalizations should be done on the combination of the two components as the independent frequency contours interact to create a third (Hoelzel & Osborne, 1986). While not widespread across animal taxa, simultaneous frequency contours have been recorded in the vocalizations of pilot whales (*Globicephala melas*), Campbell's monkey (*Cercopithecus campbelli*), emperor penguins (*Aptenodytes forsteri*) and

king penguins (*Aptenodytes patagonicus*). In the penguin species, the difference between the two frequencies is small enough that acoustic ‘beats’ appear. Measurements of these beats across individuals showed that they are individually-specific (Lengagne *et al.*, 1999a; Robisson *et al.*, 1993; Aubin *et al.*, 2000). An interaction between two vocal features allows for potentially more individual signatures, as well as increasing the signal degradation while in transit (Lengagne *et al.*, 2000). In order to accurately measure the inter-individual similarity of Campbell’s monkey vocalizations, a similarity index was used to compare call features in an integrated fashion (Lemasson *et al.*, 2003). Because the frequency modulated components of killer whale vocalizations are produced and processed simultaneously, they should also be analyzed as a single signal either with a neural network comparison approach or through another method of analysis.

In order to measure the amount of information that vocalizations may communicate, a more quantitative method should be used to weigh the relative importance of group and individual signatures within the signature system of resident killer whales. The presence of a group signature does not eliminate the possibility for an individual signature, but one could carry much more information than the other. A few studies have used information theory to evaluate the amount of information that is carried by a signal and this approach might be useful in determining the potential interaction between group and individual signatures in transmitting information (Beecher, 1989; Lengagne *et al.*, 1999b; Beecher, 1982). Similarly, killer whales have been shown to increase the duration of their calls in response to boat noise (Foote *et al.*, 2004), which suggests that

they do adjust their calls to increase the probability that their signal is received. Evaluating which features are modified to ensure the information is not lost or altered by background noise while in transit may be interesting to investigate from the perspective of information theory, as it could indicate the potential information capacity of specific call features.

Because killer whales have been shown to vary their call duration in response to increased noise levels (Foote *et al.*, 2004) and it has been suggested as a feature that may be individually specific (Hoelzel & Osborne, 1986), this may be an important feature for communication that was not explored within this study. In order to compare calls with the neural network, they had to be standardized over time. Previous comparisons of frequency contours have used the approach of this study (Deecke *et al.*, 1999; Murray *et al.*, 1998), but this standardizes the call duration without regard for the components which form the call. Ford (1987) and Miller & Bain (2000) showed that the stereotyped components can vary between different groups of animals. The method of dynamic time-warping allows call duration to be standardized while matching the components of a call through non-uniform dilation of the components independently (Buck & Tyack, 1993). This method has been used successfully to match the signature whistles of bottlenose dolphins (Buck & Tyack, 1993), as well as classify the vocalizations of transient killer whales (Deecke, 2003). Dynamic time-warping should also be applied to resident vocalizations for a more precise measurement of call similarity.

Perceptual studies in the field

Once potential individually specific features have been demonstrated in isolated vocalizations with spectral analyses, studies should be performed to assess if these structural features are perceivable within the natural habitat. As with the study of frequency components in penguins (Lengagne *et al.*, 2000), playbacks of separate components to individuals and groups involved in different behaviors would help to determine the function of each component and possibly the function of the calls. The best way to observe any response of the recipient would be to perform playbacks to an animal tagged with a detachable recording device. Similarly, playbacks of artificial modified calls would be useful in analyzing if differences in call types are perceivable by whales and/or if they elicit a particular response.

Just as the analysis of spectral features had previously only been carried out between groups, calling behavior between individuals is also relatively poorly understood. Studies of interactions between vocalizing individuals would increase the knowledge of how vocalizations function in killer whale groups. Measuring the level of similarity within vocal exchanges would explore the possibility for vocal matching on a close temporal scale. Many of the species that vocalize in matched call type bouts have also been shown to match structural features (Janik, 2000). Following the sequence of vocalizing individuals and looking for consistency in the order in which individuals vocalize could provide information about the existence of dominance hierarchies with killer whale matriline. Additionally, the sex of individuals or the relationship within

matrilineal units would be a useful subject of further analysis. More recordings of different individuals would be necessary in order to understand the extent to which sex or the proximity of genetic relatedness influences vocal similarity. The high proportion of calls from male individuals, as well as individuals of the same generation, could have produced a level of vocal similarity that was not representative of vocal similarity between all individuals.

Social structure in other cetaceans

This study has demonstrated the effectiveness of the beamforming technique in a species with individuals that rarely separate. Given the lack of knowledge about many cetacean species, especially those that are highly dependent on group living, recording vocalizations from individuals might yield a large amount of information about individual roles and interactions. The connection between vocal similarity and social structure would explain a great deal about interactions between individuals. For example, recording the calls of individual transient whales might generate an interesting intermediary example between the individual-based society of bottlenose dolphins and that of resident killer whales. Looking at the similarity levels between non-signature vocalizations of more 'resident' groups of bottlenose dolphins allows an interesting comparison, as it appears that group-specific signatures do exist in this species but are less important than those indicating individual identity (Ding et al., 1995). Just as the calls of resident killer whales may signal both group identity and individual identity, species of dolphin that have long-term associations with

specific individuals and live in areas with abundant resources may use a certain group-specific vocalization.

Conclusion

As all of these questions show, the ability to record sounds from particular cetaceans is a huge step forward and has opened up the opportunity to finally determine the full range of inter-individual vocal variation in the social interactions between individual cetaceans. Within a group-living species, the complex structure of stereotyped killer whale calls allows the transfer of large amounts of information in a very compact fashion. These calls not only pinpoint group members and coordinate their movement while foraging as a group but also contain finer scale indications for individual identification. Although the exact features that encode individual vocal signatures were not isolated in this study, the ability to distinguish between individual killer whales using only the frequency modulations of their calls has not previously been documented across several groups of individuals. This study has revealed that the strong correlation between acoustic similarity and amount of social association extends to the level of the individual, as individuals rely strongly on the use of group-specific signatures. However, there appears to be some selection for maintaining the identification of individuals within family groups so the possibility exists that individual social bonds are also important to resident killer whales.

REFERENCES

- Altmann, J. **1974**. Observational study of behavior: sampling methods. *Behaviour*, **49**, 227-267.
- Amos, B., Schlötterer, C. & Tautz, D. **1993**. Social structure of pilot whales revealed by analytical DNA profiling. *Science*, **260**, 670-671.
- Arnborn, T., Papastavrou, V., Weilgart, L. S. & Whitehead, H. **1987**. Sperm whales react to an attack by killer whales. *Journal of Mammalogy*, **68**, 450-453.
- Arnborn, T. & Whitehead, H. **1989**. Observations on the composition and behaviour of groups of female sperm whales near the Galápagos Islands. *Canadian Journal of Zoology*, **67**, 1-7.
- Aubin, T., Jouventin, P. & Hildebrand, C. **2000**. Penguins use the two-voice system to recognize each other. *Proceedings of the Royal Society of London Series B*, **267**, 1081-1087.
- Bain, D. E. **1986**. Acoustic behavior of *Orcinus*: sequences, periodicity, behavioral correlates and an automated technique for call classification. In: *Behavioral Biology of Killer Whales* (Ed. by Kirkevold B.C., L., J.S.), pp. 335-371. New York: Alan R Liss. Inc.
- Baird, R. W. **2000**. The Killer Whale: Foraging Specializations and Group Hunting. In: *Cetacean Societies: Field Studies of Dolphins and Whales* (Ed. by Mann, J., Connor, R. C., Tyack, P. L. & Whitehead, H.), pp. 127-153. Chicago, IL: University of Chicago Press.
- Baird, R. W. & Dill, L. M. **1996**. Ecological and social determinants of group size in transient killer whales. *Behavioral Ecology*, **7**, 408-416.
- Baird, R. W. & Whitehead, H. **2000**. Social organization of mammal-eating killer whales: group stability and dispersal patterns. *Canadian Journal of Zoology*, **78**, 2096-2105.
- Balcomb, K. C., Boran, J. R. & Heimlich, S. L. **1982**. Killer whales in greater Puget Sound. *Reports of the International Whaling Commission (special issue)*, **32**, 681-685.

- Balcombe, J. P. & McCracken, G. F. **1992**. Vocal recognition in Mexican free-tailed bats - do pups recognize mothers? *Animal Behaviour*, **43**, 79-87.
- Barrett-Lennard, L. **2000**. Population structure and mating patterns of killer whale populations in the Northeastern Pacific, as revealed by DNA analysis. In: *Zoology*, pp. 97. Vancouver BC: University of British Columbia.
- Barrett-Lennard, L. G., Ford, J. K. B. & Heise, K. A. **1996**. The mixed blessing of echolocation: differences in sonar use by fish-eating and mammal-eating killer whales. *Animal Behaviour*, **51**, 553-565.
- Bartlett, P. & Slater, P. J. B. **1999**. The effect of new recruits on the flock specific call of budgerigars (*Melopsittacus undulatus*). *Ethology Ecology and Evolution*, **11**, 139-147.
- Beecher, M. D. **1982**. Signature systems and kin recognition. *American Zoologist*, **22**, 477-490.
- Beecher, M. D. **1989**. Signaling systems for individual recognition: an information theory approach. *Animal Behaviour*, **38**, 248-261.
- Beecher, M. D., Stoddard, P. K., Campbell, S. E. & Horning, C. L. **1996**. Repertoire matching between neighbouring song sparrows. *Animal Behaviour*, **51**, 917-923.
- Bigg, M. A. **1982**. An assessment of killer whale (*Orcinus orca*) stocks off Vancouver Island, British Columbia. In: *Reports of the International Whaling Commission (Special Issue)*, pp. 655-666.
- Boran, J. R. & Heimlich, S. L. **1999**. Social learning in cetaceans: hunting, hearing and hierarchies. In: *Mammalian Social Learning: Comparative and Ecological Perspectives* (Ed. by Box, H. O. & Gibson, K. R.), pp. 282-307. Cambridge: Cambridge University Press.
- Boughman, J. W. **1997**. Greater spear-nosed bats give group-distinctive calls. *Behavioral Ecology and Sociobiology*, **40**, 61-70.
- Boughman, J. W. & Wilkinson, G. S. **1998**. Greater spear-nosed bats discriminate group mates by vocalizations. *Animal Behaviour*, **55**, 1717-1732.

- Buck, J. R. & Tyack, P. L. **1993**. A quantitative measure of similarity for *Tursiops truncatus* signature whistles. *Journal of the Acoustical Society of America*, **94**, 2497-2506.
- Caldwell, M. C. & Caldwell, D. k. **1965**. Individualized whistle contours in bottlenosed dolphins, *Tursiops truncatus*. *Nature (London)*, **207**, 434-435.
- Charrier, I., Mathevon, N. & Jouventin, P. **2002**. How does a fur seal mother recognize the voice of her pup? An experiemental sudy of *Arctocephalus tropicalis*. *Journal of Experimental Biology*, **205**, 603-612.
- Christal, J. & Whitehead, H. **2001**. Social affiliations within sperm whale (*Physeter macrocephalus*) long-term groups. *Ethology*, **107**, 323-340.
- Clark, C. W. & Clark, J. M. **1980**. Sound playback experiments with southern right whales (*Eubalaena australis*). *Science*, **207**, 663-665.
- Cleator, H. J., Stirling, I. & Smith, T. G. **1989**. Underwater Vocalizations of the Bearded Seal (Erignathus- Barbatus). *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **67**, 1900-1910.
- Connor, R. C., Heithaus, M. R. & Barre, L. M. **2001**. Complex social structure, alliance stability and mating access in a bottlenose dolphin 'super-alliance'. *Proceedings of the Royal Society of London Series B*, **268**, 263-267.
- Connor, R. C., Read, A. J. & Wrangham, R. W. **2000a**. Male reproductive strategies and social bonds. In: *Cetacean Societies: Field Studies of Dolphins and Whales* (Ed. by Mann, J., Connor, R. C., Tyack, P. L. & Whitehead, H.), pp. 247-269. Chicago: University of Chicago Press.
- Connor, R. C., Smolker, R. A. & Richards, A. F. **1992**. 2 levels of alliance formation among male bottle-nosed dolphins (*Tursiops* sp). *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 987-990.
- Connor, R. C., Wells, R. S., Mann, J. & Read, A. J. **2000b**. The Bottlenose Dolphin: Social Relationships in a Fission-Fusion Society. In: *Cetacean Societies: Field Studies of Dolphins and Whales* (Ed. by Mann, J., Connor, R. C., Tyack, P. L. & Whitehead, H.), pp. 91-126. Chicago: University of Chicago Press.

- Cook, M. L. H., Sayigh, L. S., Blum, J. E. & Wells, R. S. **2004**. Signature-whistle production in undisturbed free-ranging bottlenose dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society of London Series B*, **271**, 1043-1049.
- de March, B. G. E. & Postma, L. D. **2003**. Molecular genetic stock discrimination of belugas (*Delphinapterus leucas*) hunted in eastern Hudson Bay, northern Quebec, Hudson Strait, and Sanikiluaq (Belcher Islands), Canada, and comparisons to adjacent populations. *Arctic*, **56**, 111-124.
- Deecke, V. B. **2003**. The vocal behaviour of transient killer whale (*Orcinus orca*): communicating with costly calls. PhD Thesis. School of Biology, University of St Andrews. 134 pp.
- Deecke, V. B., Ford, J. K. B. & Spong, P. **1999**. Quantifying complex patterns of bioacoustic variation: Use of a neural network to compare killer whale (*Orcinus orca*) dialects. *Journal of the Acoustical Society of America*, **105**, 2499-2507.
- 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.
- Deecke, V. B., Slater, P. J. B. & Ford, J. K. B. **2002**. Selective habituation shapes acoustic predator recognition in harbour seals. *Nature*, **420**, 171-173.
- Ding, W., Würsig, B. & Evans, W. E. 1995. Comparisons of whistles between populations of bottlenose dolphins. *Aquatic Mammals*, **25**, 508-517.
- Domenici, P., Batty, R. S., Simila, T. & Ogam, E. **2000**. Killer whales (*Orcinus orca*) feeding on schooling herring (*Clupea harengus*) using underwater tail-slaps: kinematic analyses of field observations. *Journal of Experimental Biology*, **203**, 283-294.
- Falls, J. B. **1982**. Individual Recognition by Sounds in Birds. In: *Acoustic Communication in Birds* (Ed. by Kroodsma, D. E., Miller, E.H., Ouellet, H.), pp. 237-278. London: Academic Press, Inc.
- Farabaugh, S. M., Brown, E. D. & Veltman, C. J. **1988**. Song sharing in a group-living songbird, the Australian magpie. Part II. Vocal sharing between

- territorial neighbors, within and between geographic regions and between sexes. *Behaviour*, **104**, 105-125.
- Feekes, F. **1982**. Song mimesis within colonies of Caciús-C Cela (Icteridae, Aves)- A colonial password. *Zeitschrift für Tierpsychologie*, **58**, 119-152.
- Foot, A. D., Osborne, R. W. & Hoelzel, A. R. **2004**. Whale-call response to masking boat noise. *Nature*, **428**, 910.
- Ford, J. K. B. **1987**. A Catalogue of Underwater Calls Produced by Killer Whales (*Orcinus orca*) in British Columbia. pp. 165. Nanaimo, British Columbia: Department of Fisheries and Oceans, Fisheries Research Branch.
- Ford, J. K. B. **1989**. Acoustic behaviour of resident killer whale (*Orcinus orca*) off Vancouver Island, British Columbia. *Canadian Journal of Zoology*, **67**, 727-745.
- 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.
- Ford, J. K. B., Ellis, G., Barrett-Lennard, L., Morton, A. B., Palm, R. S. & Balcomb, K. C. **1998**. Dietary specialization in two sympatric populations of killer whales (*Orcinus orca*) in coastal British Columbia and adjacent waters. *Canadian Journal of Zoology*, **76**, 1456-1571.
- Ford, J. K. B., Ellis, G. M. & Balcomb, K. C. **2000**. *Killer Whales; The natural history and genealogy of Orcinus orca in British Columbia and Washington State*. Vancouver: UBC Press.
- Gelfand, D. L. & McCracken, G. F. **1986**. Individual variation in the isolation calls of Mexican free-tailed bat pups (*Tadarida-Brasiliensis Mexicana*). *Animal Behaviour*, **34**, 1078-1086.
- Gowans, S., Whitehead, H. & Hooker, S. K. **2001**. Social organization in northern bottlenose whales, *Hyperoodon ampullatus*: not driven by deep-water foraging? *Animal Behaviour*, **62**, 369-377.
- Green, S. **1975**. Dialects in Japanese monkeys: vocal learning and cultural transmission of local specific behavior. *Zeitschrift für Tierpsychologie*, **38**, 304-314.

- Guinet, C. **1991**. Intentional stranding apprenticeship and social play in killer whales (*Orcinus orca*). *Canadian Journal of Zoology*, **69**, 2712-2716.
- Guinet, C. & Bouvier, J. **1995**. Development of intentional stranding hunting techniques in killer whale (*Orcinus orca*) calves at Crozet Archipelago. *Canadian Journal of Zoology*, **73**, 27-33.
- Hauser, M. D. **1997**. *The Evolution of Communication*. Cambridge, MA: MIT Press.
- Haykin, S. **1994**. *Neural Networks: A Comprehensive Foundation*. New York: Macmillan College Publishing Company.
- Heide-Jorgensen, M. P., Richard, P., Dietz, R., Laidre, K. L., Orr, J. & Schmidt, H. C. **2003**. An estimate of the fraction of belugas (*Delphinapterus leucas*) in the Canadian high Arctic that winter in West Greenland. *Polar Biology*, **26**, 318-326.
- Heimlich-Boran, J. R. **1986a**. Fishery correlations with the occurrence of killer whales in greater Puget Sound. In: *Behavioral Biology of Killer Whales* (Ed. by Kirkevoold B.C., L., J.S.), pp. 113-131. New York: Alan R. Liss, Inc.
- Heimlich-Boran, S. L. **1986b**. Cohesive relationships among Puget Sound killer whales. In: *Behavioral Biology of Killer Whales* (Ed. by Kirkevoold B.C., L., J.S.), pp. 251-284. New York: Alan R. Liss, Inc.
- Hinton, G. E. **1992**. How neural networks learn from experience. *Scientific American*, 105-109.
- Hoelzel, A. R. **1993**. Foraging behavior and social group dynamics in Puget Sound killer whales. *Animal Behaviour*, **45**, 581-591.
- Hoelzel, A. R., Dahlheim, M. & Stern, S. J. **1998**. Low genetic variation among killer whales (*Orcinus orca*) in the eastern North Pacific and genetic differentiation between foraging specialists. *Journal of Heredity*, **89**, 121-128.
- Hoelzel, A. R. & Dover, G. A. **1991**. Genetic differentiation between sympatric killer whale populations. *Heredity*, **66**, 191-195.

- Hoelzel, A. R., Natoli, A., Dahlheim, M. E., Olavarria, C., Baird, R. W. & Black, N. A. 2002. Low worldwide genetic diversity in the killer whale (*Orcinus orca*): implications for demographic history. *Proceedings of the Royal Society of London Series B*, **269**, 1467-1473.
- Hoelzel, A. R. & Osborne, R. W. 1986. Killer whale call characteristics: Implications for cooperative foraging strategies. In: *Behavioral Biology of Killer Whales* (Ed. by Kirkevoold B.C., L., J.S.), pp. 373-403. New York: Alan R. Liss, Inc.
- Hopp, S. L., Jablonski, P. & Brown, J. L. 2001. Recognition of group membership by voice in Mexican jays, *Aphelocoma ultramarina*. *Animal Behaviour*, **62**, 297-303.
- Insley, S. J. 1992. Mother/offspring separation and acoustic stereotypy: A comparison of call morphology in two species of pinnipeds. *Behaviour*, **120**, 103-121.
- Janik, V. M. 2000a. Source levels and the estimated active space of bottlenose dolphin (*Tursiops truncatus*) whistles in the Moray Firth, Scotland. *Journal of Comparative Physiology*, **186**, 673-681.
- Janik, V. M. 2000b. Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science*, **289**, 1355-1357.
- Janik, V. M., Dehnhardt, G. & Todt, D. 1994. Signature whistle variations in a bottlenosed dolphin, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, **35**, 243-248.
- Janik, V. M. & Slater, P. J. B. 1998. Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls. *Animal Behaviour*, **56**, 829-838.
- Johnson, M. P. & Tyack, P. L. 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *Ieee Journal of Oceanic Engineering*, **28**, 3-12.
- Jouventin, P., Aubin, T. & Lengagne, T. 1999. Finding a parent in a king penguin colony: the acoustic system of individual recognition. *Animal Behaviour*, **57**, 1175-1183.

- Koene, P. **1996**. Temporal structure of red jungle fowl crow sequences: Single-case analysis. *Behavioural Processes*, **38**, 193-202.
- Krützen, M., Barré, L. M., Connor, R. C., Mann, J. & Sherwin, W. B. **2004**. 'O father: where art thou?' - Paternity assessment in an open fission-fusion society of wild bottlenose dolphins (*Tursiops* sp) in Shark Bay, Australia. *Molecular Ecology*, **13**, 1975-1990.
- Lemasson, A., Gautier, J.-P. & Hausberger, M. **2003**. Vocal similarities and social bonds in Campbell's monkey (*Cercopithecus campbelli*). *C. R. Biologies*, **326**, 1185-1193.
- Lemon, R. E. & Chatfield, C. **1971**. Organization of song in Cardinals. *Animal Behaviour*, **19**, 1-17.
- Krause, J. & Ruxton, G. D. **2002**. *Living in Groups*. Oxford: Oxford University Press.
- Lengagne, T., Aubin, T., Jouventin, P. & Lauga, J. **1999a**. Acoustic communication in a king penguin colony: importance of bird location within the colony and of the body position of the listener. *Polar Biology*, **21**, 262-268.
- Lengagne, T., Aubin, T., Lauga, J. & Jouventin, P. **1999b**. How do king penguins (*Aptenodytes patagonicus*) apply the mathematical theory of information to communicate in windy conditions? *Proceedings of the Royal Society of London Series B*, **266**, 1623-1628.
- Lengagne, T., Aubin, T., Jouventin, P. & Lauga, J. **2000**. Perceptual salience of individually distinctive features in the calls of adult king penguins. *Journal of the Acoustical Society of America*, **107**, 508-516.
- Lengagne, T., Lauga, J. & Aubin, T. **2001**. Intra-syllabic acoustic signatures used by the king penguin in parent-chick recognition: An experimental approach. *Journal of Experimental Biology*, **204**, 663-672.
- Lopez, J. C. & Lopez, D. **1985**. Killer whales (*Orcinus orca*) of Patagonia, and their behaviour of intentionally stranding while hunting nearshore. *Journal of Mammalogy*, **66**, 181-183.
- Madsen, P. T., Wahlberg, M. & Möhl, B. **2002**. Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: implications for

- echolocation and communication. *Behavioral Ecology and Sociobiology*, **53**, 31-41.
- McCowan, B. 1995. A new quantitative technique for categorizing whistles using simulated signals and whistles from captive bottle-nosed dolphins (Delphinidae, *Tursiops-truncatus*). *Ethology*, **100**, 177-193.
- McCracken, G. F. 1993. Locational memory and female pup reunions in Mexican free-tailed bat maternity colonies. *Animal Behaviour*, **45**, 811-813.
- Medvin, M. B., Stoddard, P. K. & Beecher, M. D. 1992. Signals for parent-offspring recognition - strong sib-sib call similarity in cliff swallows but not barn swallows. *Ethology*, **90**, 17-28.
- Miller, P. J. O. & Tyack, P. L. 1998. A small towed beamforming array to identify vocalizing resident killer whales (*Orcinus orca*) concurrent with focal behavioral observations. *Deep Sea Research II*, **45**, 1389-1405.
- Miller, P. J. O. & Bain, D. E. 2000. Within-pod variation in the sound production of a pod of killer whales, *Orcinus orca*. *Animal Behaviour*, **60**, 617-628.
- Miller, P. J. O. 2002. Mixed-directionality of killer whale stereotyped calls: a direction of movement cue? *Behavioral Ecology and Sociobiology*, **52**, 262-270.
- Miller, P. J. O., Johnson, M. P., Tyack, P. L. & Terray, E. A. 2004a. Swimming gaits, passive drag and buoyancy of diving sperm whales *Physeter macrocephalus*. *Journal of Experimental Biology*, **207**, 1953-1967.
- Miller, P. J. O., Shapiro, A. D., Tyack, P. L. & Solow, A. R. 2004b. Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*. *Animal Behaviour*, **67**, 1099-1107.
- Mitani, J. C., Hasegawa, T., Gros-louis, J., Marler, P. & Byrne, R. 1992. Dialects in Wild Chimpanzees. *American Journal of Primatology*, **27**, 233-243.
- Mitani, J. C., Hunley, K. L. & Murdoch, M. E. 1999. Geographic variation in the calls of wild chimpanzees: A reassessment. *American Journal of Primatology*, **47**, 133-151.

- Möller, L. M. & Beheregaray, L. B. **2004**. Genetic evidence for sex-biased dispersal in resident bottlenose dolphins (*Tursiops aduncus*). *Molecular Ecology*, **13**, 1607-1612.
- Möller, L. M., Beheregaray, L. B., Harcourt, R. G. & Krützen, M. **2001**. Alliance membership and kinship in wild male bottlenose dolphins (*Tursiops aduncus*) of southeastern Australia. *Proceedings of the Royal Society of London Series B*, **268**, 1941-1947.
- Morton, A. B., Gale, J. C. & Prince, R. C. **1986**. Sound and behavioral correlations in captive *Orcinus orca*. In: *Behavioral Biology of Killer Whales* (Ed. by Kirkevoold B.C., L., J.S.), pp. 303-333. New York: Alan R. Liss, Inc.
- Mundinger, P. **1970**. Vocal imitation and individual recognition of finch calls. *Science*, **168**, 480-482.
- Murray, S. O., Mercado, E. & Roitblat, H. L. **1998**. The neural network classification of false killer whale (*Pseudorca crassidens*) vocalizations. *Journal of the Acoustical Society of America*, **104**, 3626-3633.
- Nichol, L. M. & Shackleton, D. M. **1996**. Seasonal movements and foraging behaviour of northern resident killer whales (*Orcinus orca*) in relation to the inshore distribution of salmon (*Oncorhynchus* spp) in British Columbia. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **74**, 983-991.
- Nowacek, D. P., Johnson, M. P. & Tyack, P. L. **2004**. North Atlantic right whales (*Eubalaena glacialis*) ignore ships but respond to alerting stimuli. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **271**, 227-231.
- Olesiuk, P. K., Bigg, M. A. & Ellis, G. **1990**. Life history and population dynamics of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia. *Reports of the International Whaling Commission (special issue)*, **12**.

- Osborne, R. W. **1986**. A behavioral budget of Puget Sound killer whales. In: *Behavioral Biology of Killer Whales* (Ed. by Kirkevold B.C., L., J.S.), pp. 211-249. New York: Alan R. Liss, Inc.
- Ottensmeyer, C. A. & Whitehead, H. **2003**. Behavioural evidence for social units in long-finned pilot whales. *Canadian Journal of Zoology*, **81**, 1327-1338.
- Payne, R. S. & McVay, S. **1971**. Songs of humpback whales. *Science*, **173**, 585-597.
- Phillips, A. V. & Stirling, I. **2000**. Vocal individuality in mother and pup South American fur seals, *Arctocephalus australis*. *Marine Mammal Science*, **16**, 593-616.
- Pitman, R. L., Balance, L. T., Mesnick, S. L. & Chivers, S. **2001**. Killer whale predation on sperm whales: Observations and implications. *Marine Mammal Science*, **17**, 494-507.
- Price, J. J. **1999**. Recognition of family-specific calls in stripe-backed wrens. *Animal Behaviour*, **57**, 483-492.
- Price, J. J. **2003**. Communication with shared call repertoires in the cooperatively breeding Stripe-backed Wren. *Journal of Field Ornithology*, **74**, 166-171.
- Ralston, J. V. & Herman, L. M. **1995**. Perception and generalization of frequency contours by a bottlenose dolphin (*Tursiops truncatus*). *Journal of Comparative Psychology*, **109**, 268-277.
- Rendell, L. E. & Whitehead, H. **2003**. Vocal clans in sperm whales (*Physeter macrocephalus*). *Proceedings of the Royal Society of London Series B*, **270**, 225-231.
- Rendell, L. E. & Whitehead, H. **2004**. Do sperm whales share coda vocalizations? Insights into coda usage from acoustic size measurements. *Animal Behaviour*, **67**, 865-874.
- Ridgeway, S., Carder, D., Smith, R., Kamolnick, T. & Elsberry, W. **1997**. First audiogram for marine mammals in the open ocean and at depth: hearing and whistling by two white whales down to 30 atmospheres. *Journal of the Acoustical Society of America*, **101**, 3136.

- Robisson, P. **1992**. Vocalizations in *Aptenodytes* penguins: application of the two-voice theory. *Auk*, **109**, 654-658.
- Robisson, P., Aubin, T. & Bremond, J. C. **1993**. Individuality in the voice of the emperor penguin *Aptenodytes forsteri*: Adaption to a noisy environment. *Ethology*, **94**, 279-290.
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. **1986**. Learning internal representations by error propagation. In: *Parallel Distributed Processing* (Ed. by Rumelhart, D. E., McClelland, J.L., PDP Research Group), pp. 318-362. Cambridge, MA: MIT Press.
- Saulitis, E., Matkin, C., Barrett-Lennard, L., Heise, K. & Ellis, G. **2000**. Foraging strategies of sympatric killer whale (*Orcinus orca*) populations in Prince William Sound, Alaska. *Marine Mammal Science*, **16**, 94-109.
- Sayigh, L. S., Tyack, P. L., Wells, R. S., Solows, A. R., Scott, M. D. & Irvine, A. B. **1998**. Individual recognition in wild bottlenose dolphins: a field test using playback experiments. *Animal Behaviour*, **57**, 41-50.
- Scherrer, J. A. & Wilkinson, G. S. **1993**. Evening bat isolation calls provide evidence for heritable signatures. *Animal Behaviour*, **46**, 847-860.
- Searby, A., Jouventin, P. & Aubin, P. **2004**. Acoustic recognition in macaroni penguins: an original signature system. *Animal Behaviour*, **67**, 615-625.
- Similä, T. & Ugarte, F. **1993**. Surface and underwater observations of cooperatively feeding killer whales, *Orcinus orca* Linnaeus 1758, in the waters of northern Norway. *Canadian Journal of Zoology*, **71**, 1494-1499.
- Simila, T., Holst, J. C. & Christensen, I. **1996**. Occurrence and diet of killer whales in northern Norway: Seasonal patterns relative to the distribution and abundance of Norwegian spring-spawning herring. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 769-779.
- Slater, P. J. B. **1973**. Describing sequences of behavior. In: *Perspectives in Ethology* (Ed. by Bateson, P. P. G. & Klopfer, P. H.), pp. 131-153. London: Plenum Press.
- Slater, P. J. B. & Lester, N. P. **1982**. Minimising errors in splitting behaviour into bouts. *Behaviour*, **79**, 153-161.

- Smolker, R. A., Mann, J. & Smuts, B. B. **1993**. Use of signature whistles during separations and reunions by wild bottlenose dolphin mothers and infants. *Behavioral Ecology and Sociobiology*, **33**, 393-402.
- Smolker, R. A., Richards, A. F., Connor, R. C. & Pepper, J. W. **1992**. Sex differences in patterns of association among Indian Ocean bottlenose dolphins. *Behaviour*, **123**, 38-69.
- Sokal, R.R. & Rohlf, F.J. **1981**. Biometry: The Principles and Practice of Statistics in Biological Research. New York: W.H.Freeman & Co. 859 p.
- Strager, H. **1995**. Pod-specific call repertoires and compound calls of killer whales, *Orcinus orca* Linnaeus, 1758, in the waters of northern Norway. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, **73**, 1037-1047.
- Thomas, J. & Stirling, I. **1983**. Geographic variation in the underwater vocalizations of weddell seals (*Leptonychotes weddelli*) from Palmer Peninsula and McMurdo Sound, Antarctica. *Canadian Journal of Zoology*, **61**, 2203-2212.
- Thomsen, F., Franck, D. & Ford, J. K. B. **2001**. Characteristics of whistles from the acoustic repertoire of resident killer whales (*Orcinus orca*) off Vancouver Island, British Columbia. *Journal of the Acoustical Society of America*, **109**, 1240-1246.
- Thomsen, F., Franck, D. & Ford, J. K. B. **2002**. On the communicative significance of whistles in wild killer whales (*Orcinus orca*). *Naturwissenschaften*, **89**, 404-407.
- Trillmich, F. **1981**. Mutual mother-pup recognition in Galapagos fur seals and sea lions: cues used and functional significance. *Behaviour*, **78**, 21-42.
- Tyack, P. L. **1986a** Whistle repertoires of two bottlenose dolphins, *Tursiops truncatus*: mimicry of signature whistles? *Behavioral Ecology and Sociobiology*, **18**, 251-257.
- Tyack, P. L. **1986b** Population biology, social behavior and communication in whales and dolphins. *Trends in Ecology and Evolution*, **1**, 144-150.

- Tyack, P. **1997**. Development and social functions of signature whistles in bottlenose dolphins *Tursiops truncatus*. *Bioacoustics*, **8**, 21-46.
- Tyack, P. **1998**. Acoustic communication under the sea. In: *Animal Acoustic Communication: Sound Analysis and Research Methods* (Ed. by Hopp, S. L., Owren, M.J., Evans, C.S.), pp. 163-213. Berlin: Springer.
- Tyack, P. L. & Sayigh, L. S. **1997**. Vocal learning in cetaceans. In: *Social influences on vocal development* (Ed. by Snowdon, C. T. & Hausberger, M.), pp. 208-233. Cambridge: Cambridge University Press.
- Wang, C. & Seneff, S. **2000**. Robust pitch tracking for prosodic modeling in telephone speech. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. Istanbul, Turkey.
- Watkins, W. A. & Schevill, W. E. 1977. Sperm whale codas. *Journal of the Acoustical Society of America*, **62**, 1486-1490.
- Watwood, S. L., Tyack, P. L. & Wells, R. S. **2004**. Whistle sharing in paired male bottlenose dolphins, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, **55**, 531-543.
- Weilgart, L. & Whitehead, H. 1997. Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whales. *Behavioral Ecology and Sociobiology*, **40**, 277-285.
- Weilgart, L. S. & Whitehead, H. **1993**. Coda communication by sperm whale (*Physeter macrocephalus*) off the Galapagos Islands. *Canadian Journal of Zoology*, **71**, 744-752.
- Whitehead, H. & Weilgart, L. S. **1991**. Patterns of visually observable behaviour and vocalization in groups of female sperm whales. *Behaviour*, **118**, 275-296.
- Whitehead, H. **1995**. Investigating structure and temporal scale in social organizations using identified individuals. *Behavioral Ecology*, **6**, 199-208.
- Whitehead, H. **2003**. *Sperm Whales: Social Evolution in the Ocean*. Chicago, IL: University of Chicago Press.

- Whitehead, H., Gowans, S., Faucher, A. & McCarrey, S. W. 1997. Population analysis of northern bottlenose whales in the Gully, Nova Scotia. *Marine Mammal Science*, **13**, 173-185.
- 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**, 385-389.
- Wilkinson, G. S. & Boughman, J. W. 1998. Social calls coordinate foraging in greater spear-nosed bats. *Animal Behaviour*, **55**, 337-350.
- Wilkinson, G. S. & Boughman, J. W. 1999. Social influences on foraging in bats. In: *Mammalian Social Learning: Comparative and Ecological Perspectives* (Ed. by Box, H. O. & Gibson, K. R.), pp. 188-204. Cambridge: Cambridge University Press.
- Yurk, H., Barrett-Lennard, L., Ford, J. K. B. & Matkins, C. O. 2002. Cultural transmission within maternal lineages: vocal clans in resident killer whales in southern Alaska. *Animal Behaviour*, **63**, 1103-1119.
- Zar, J. H. 1996. *Biostatistical Analysis*. Upper Saddle River: Prentice-Hall.

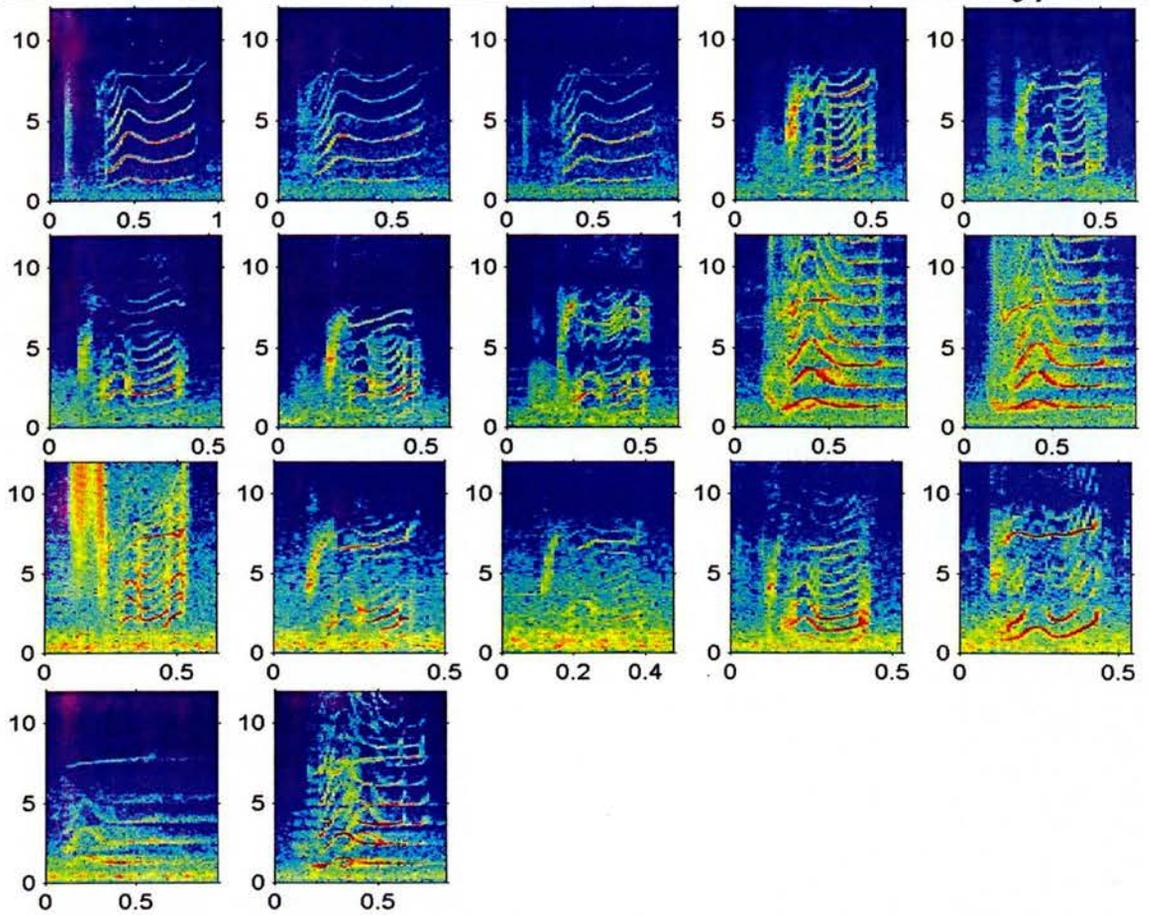
APPENDIX 1: SPECTROGRAMS OF STEREOTYPED CALLS FROM INDIVIDUAL RESIDENT KILLER WHALES

The following pages include spectrograms of all vocalizations that were included in the final analysis of call similarity. All of these calls were recorded from northern resident killer whales inhabiting the Johnstone Strait area off Vancouver Island, British Columbia during August and September 1998 and 1999. Calls are listed by (1) the individual that produced them, (2) the matriline to which that individual belongs and (3) call type. All individual whales and their matriline are identified by either an 'A' or a 'W' and all call types begin with a 'N' as they are all from members of the northern resident community. See Ford (1987) for a complete explanation of the distinguishing features of each call type. See Chapter 2 for a detailed description of the recording process.

Spectrograms used a sampling rate of 48 kHz, an FFT size of 2048 samples, a dynamic range of 50 dB and a maximum frequency limit of 11 kHz and were created using custom MATLAB software. Additional custom programs allowed calls of specific individuals to be selected according to call type. All spectrograms show time in seconds on the x axis and frequency in kilohertz on the y axis.

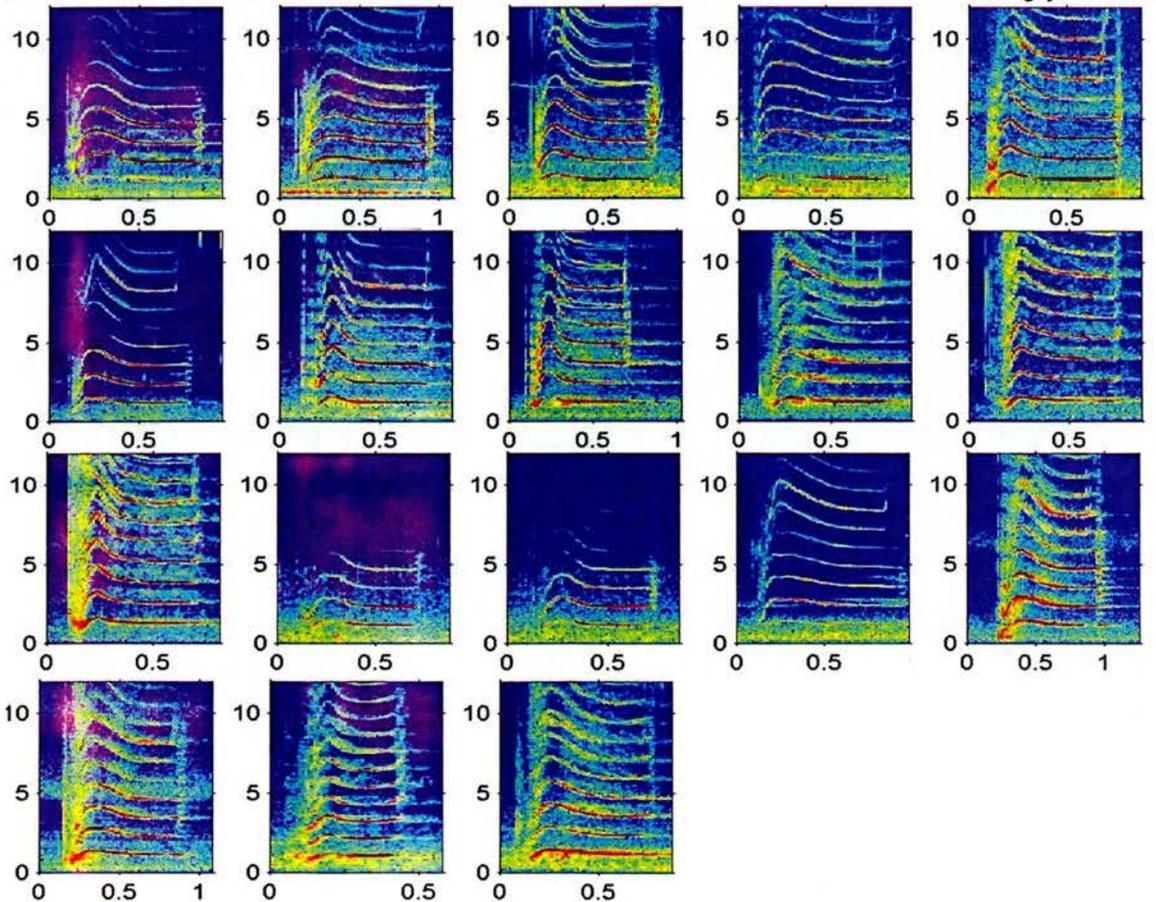
A32 – A36 Matriline

Call Type N2



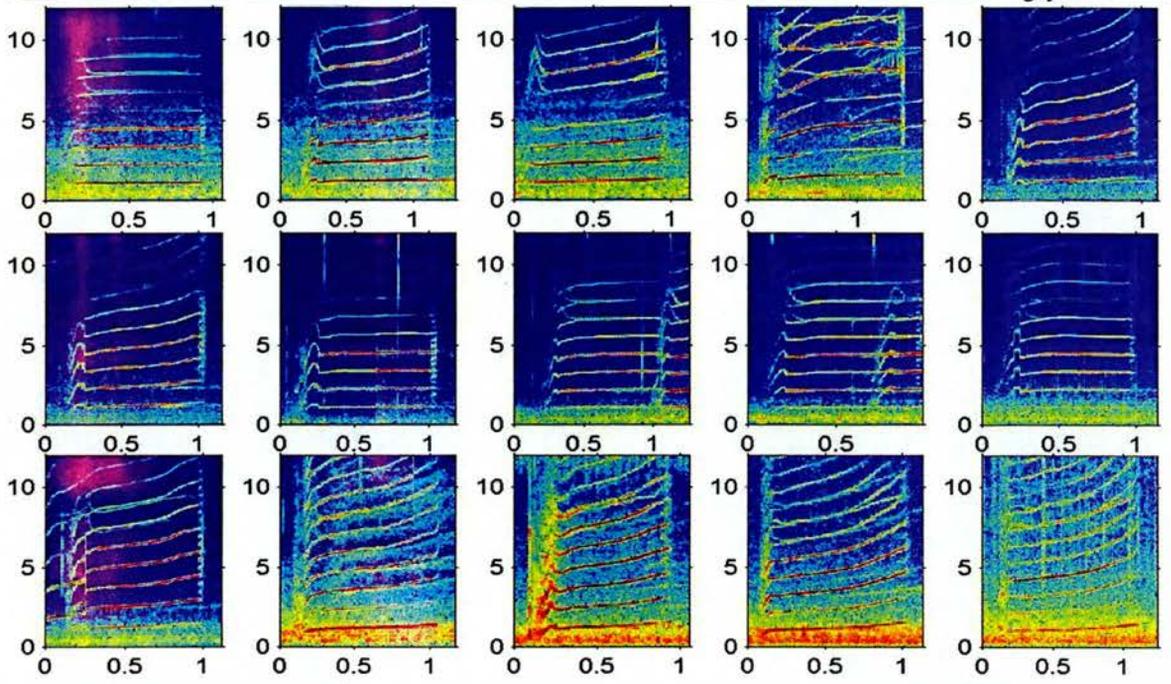
A32 – A36 Matriline

Call Type N4



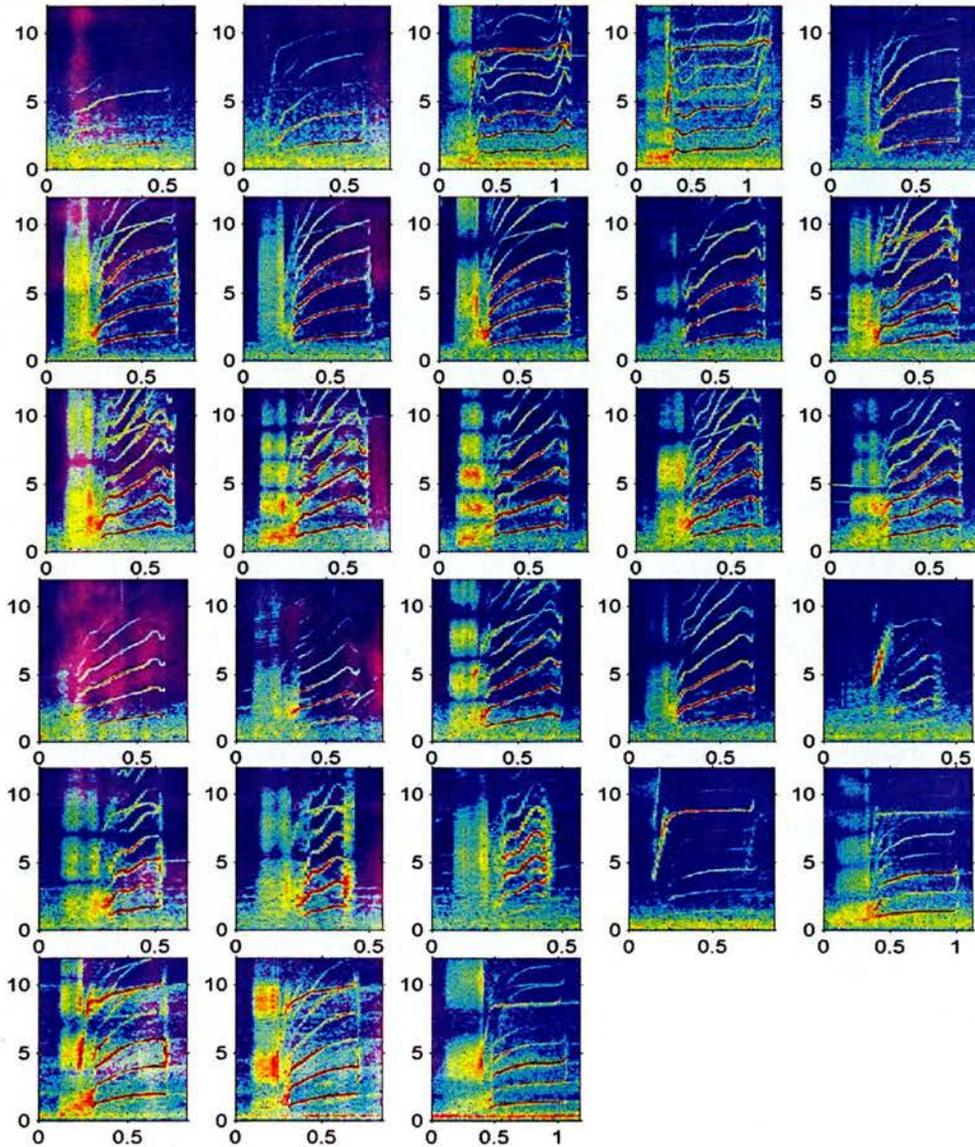
A32 – Continued

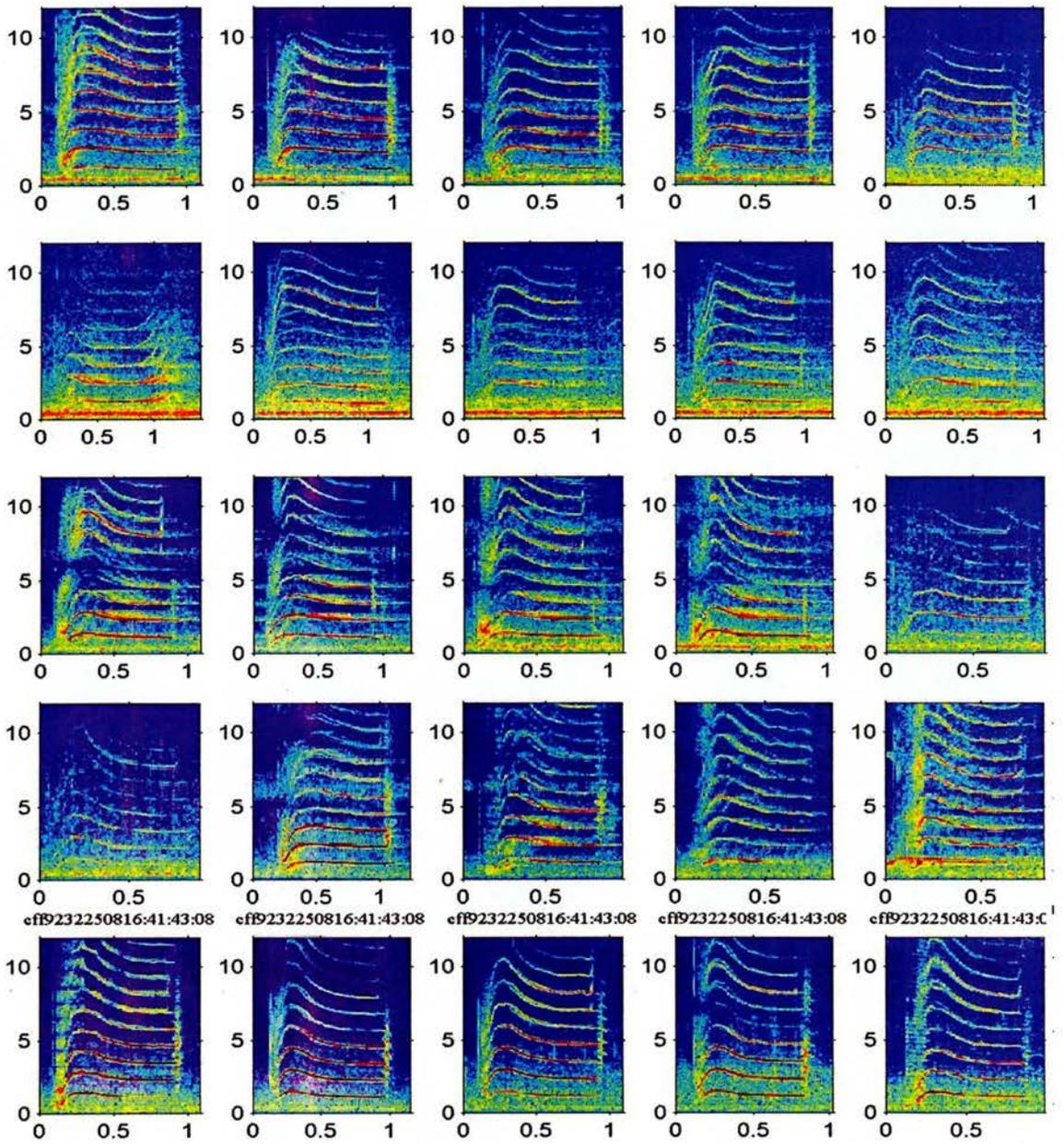
Call Type N5

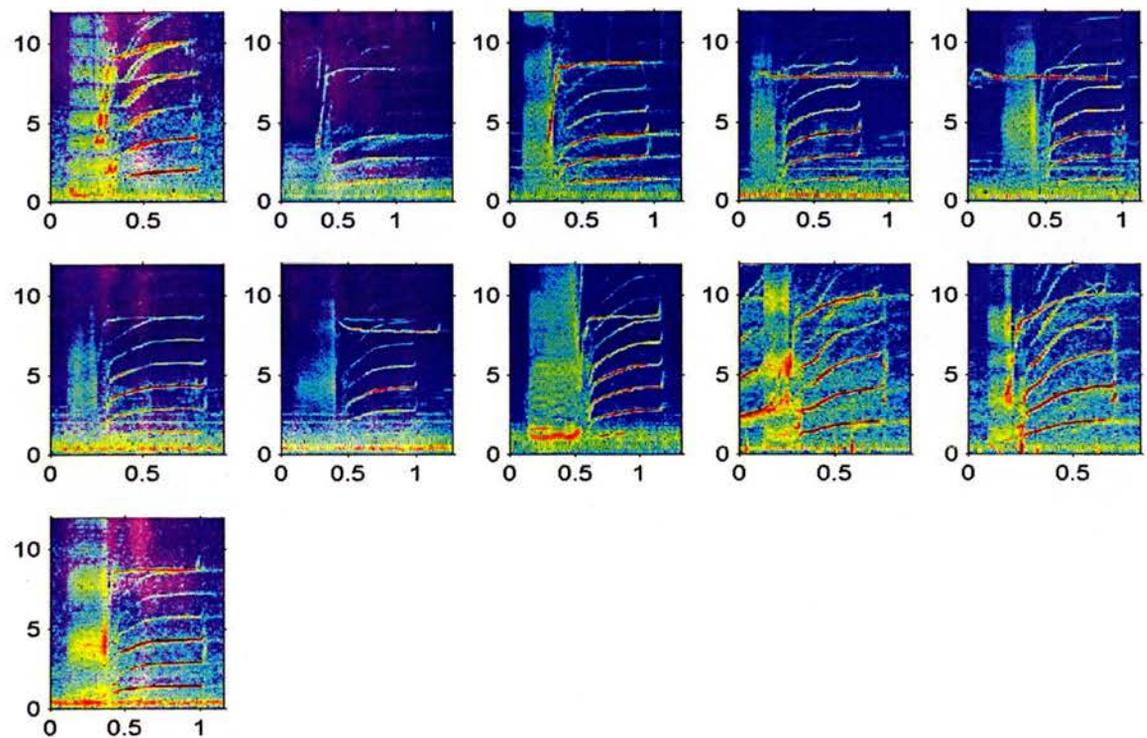
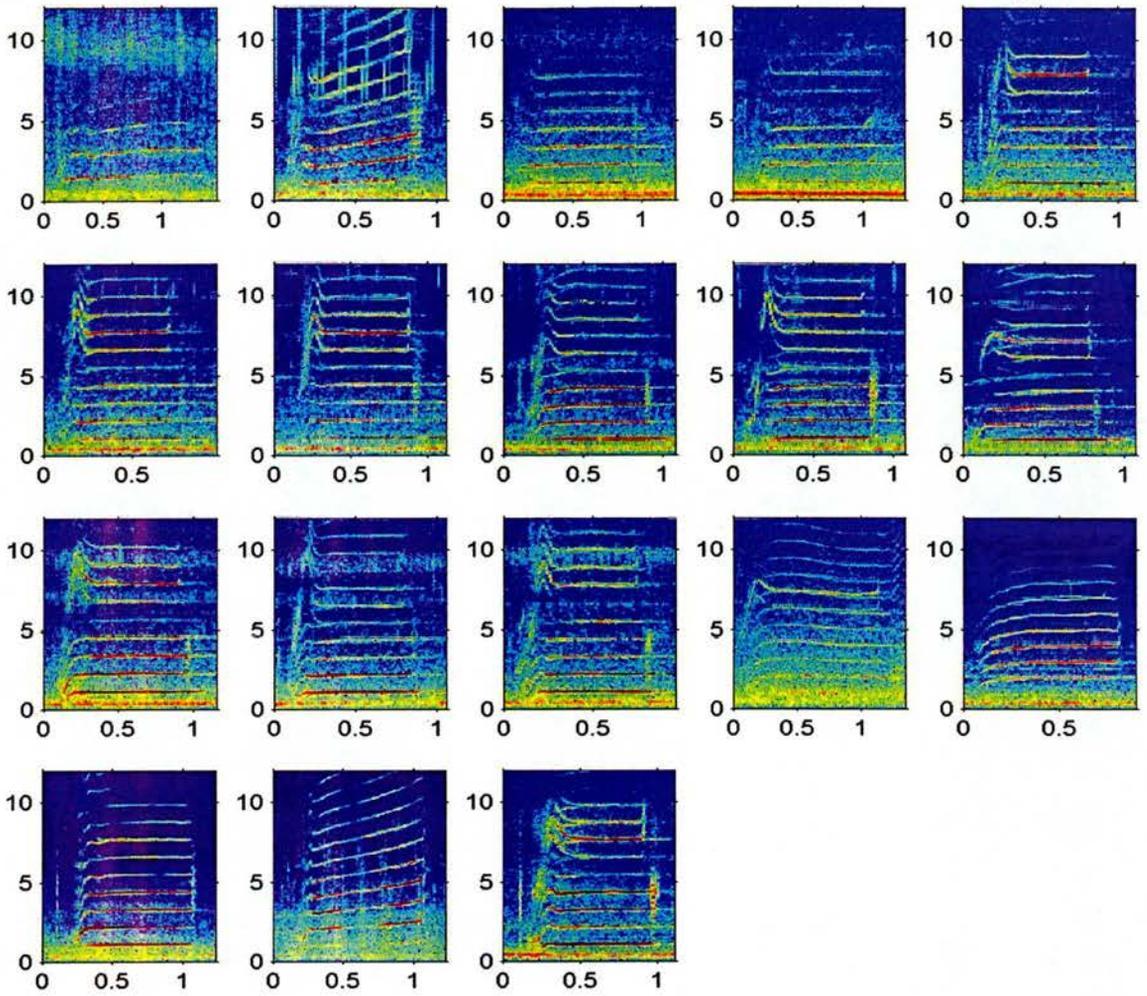


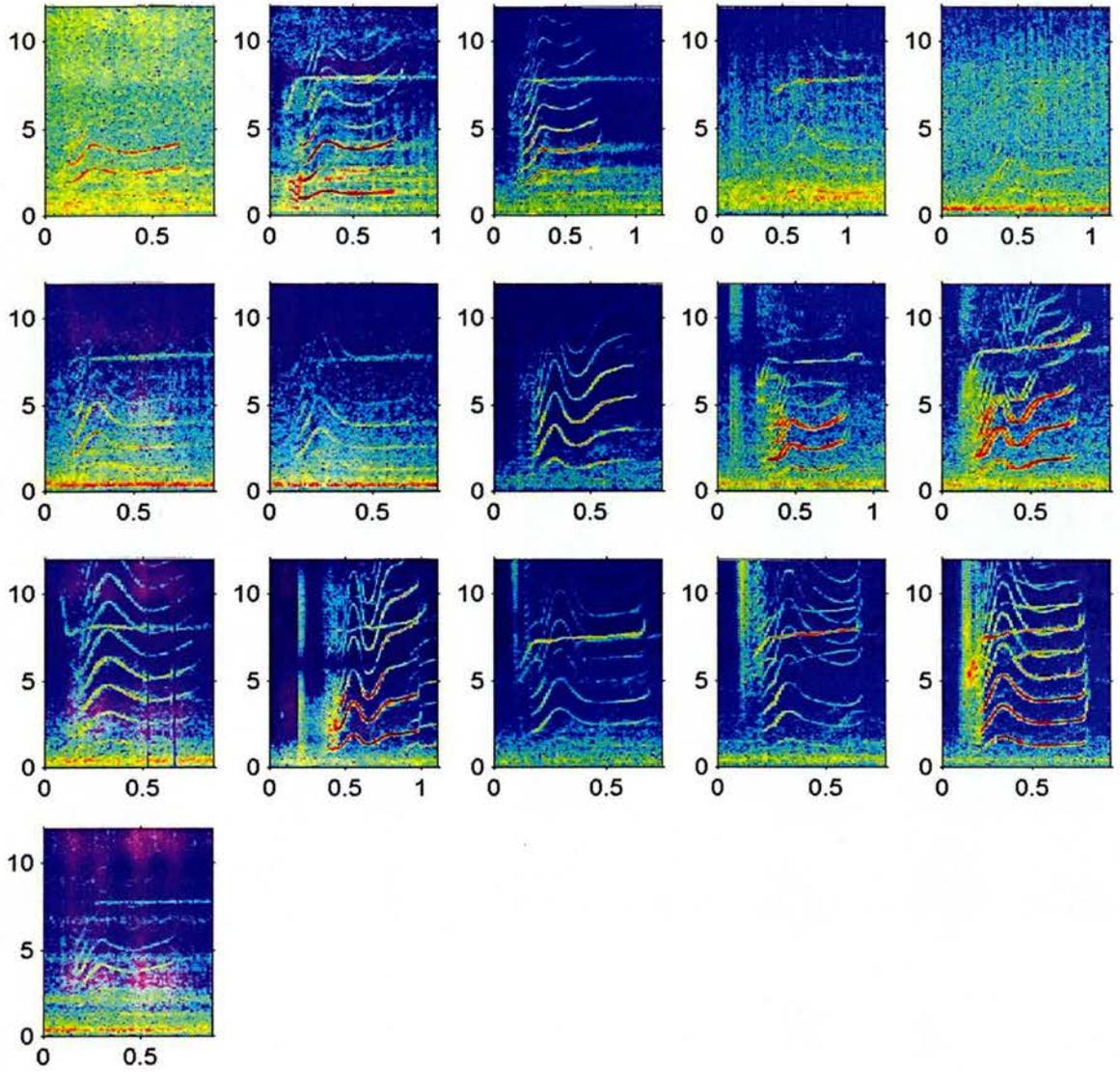
A32 – Continued

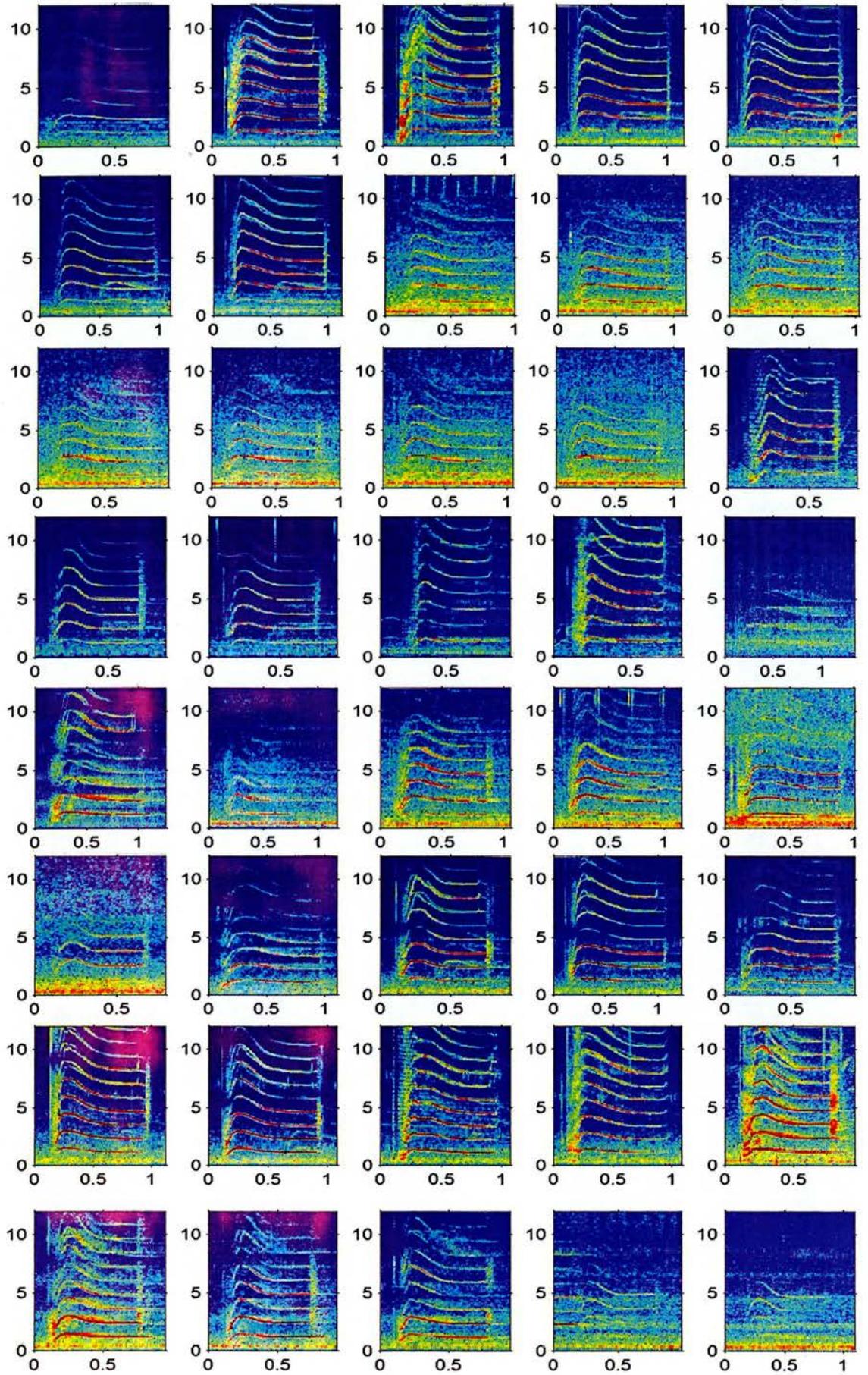
Call Type N9

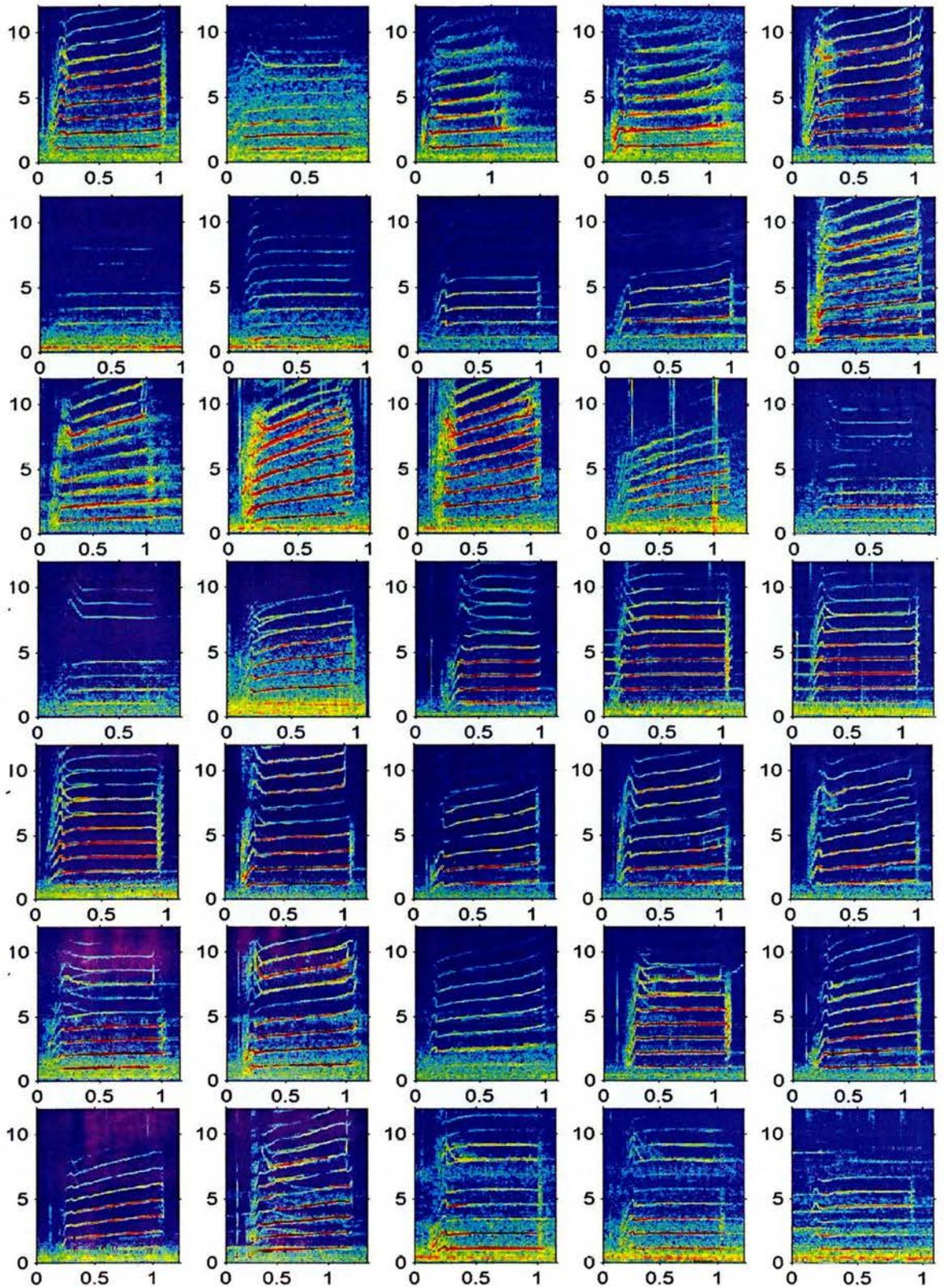


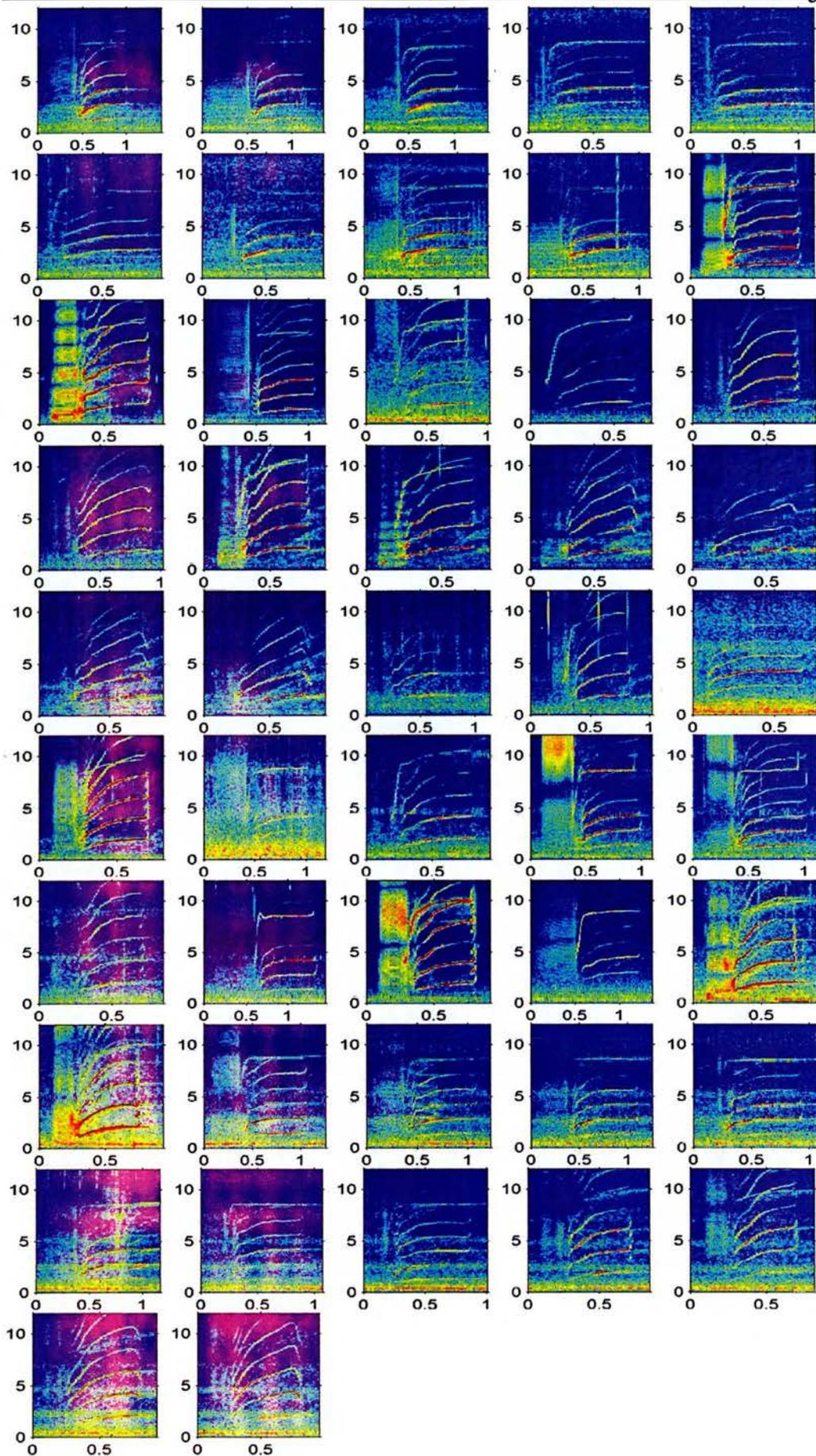






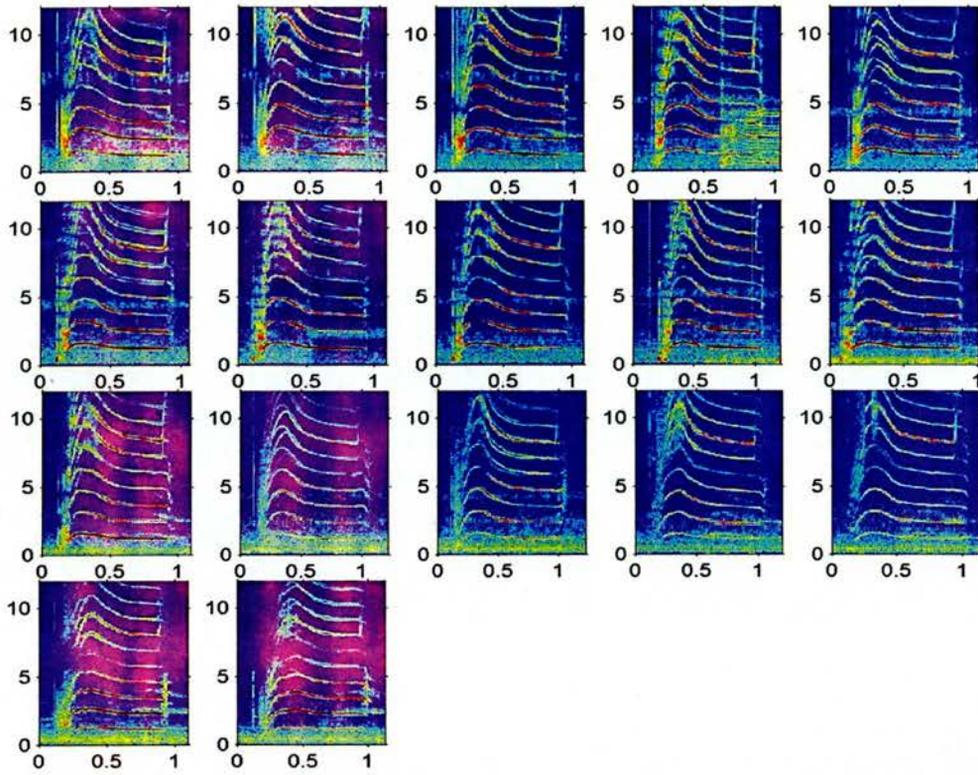






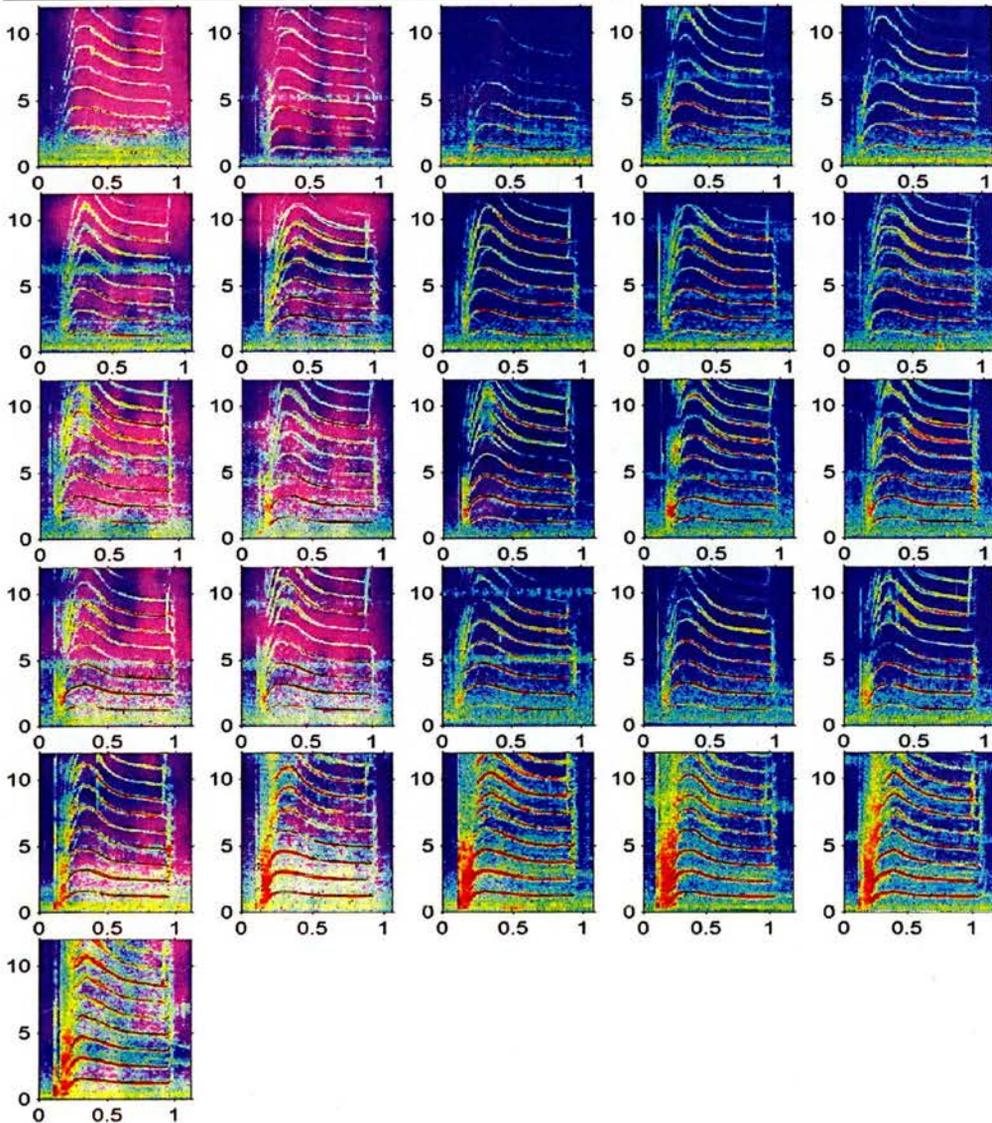
A6 – Matriline A30

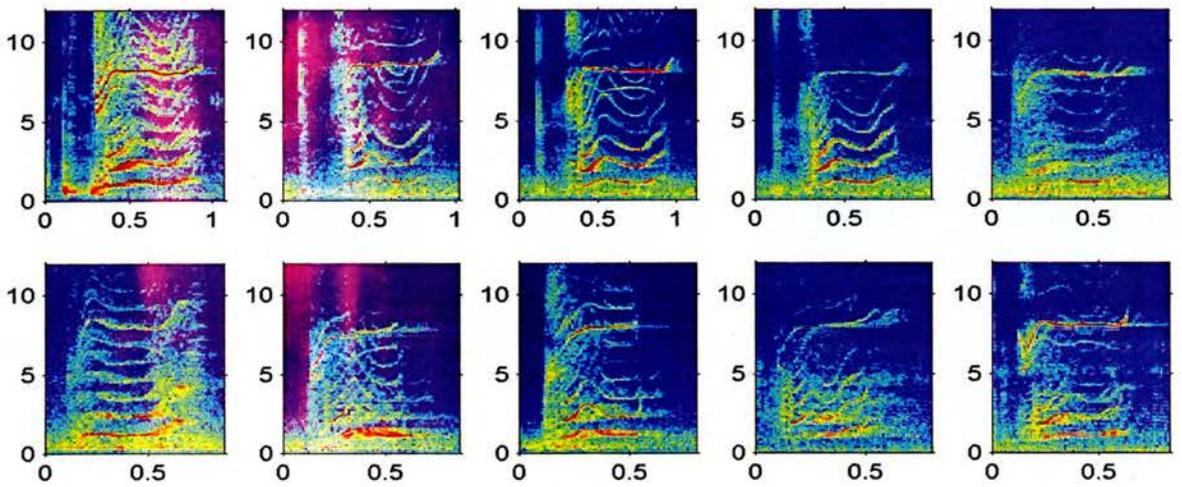
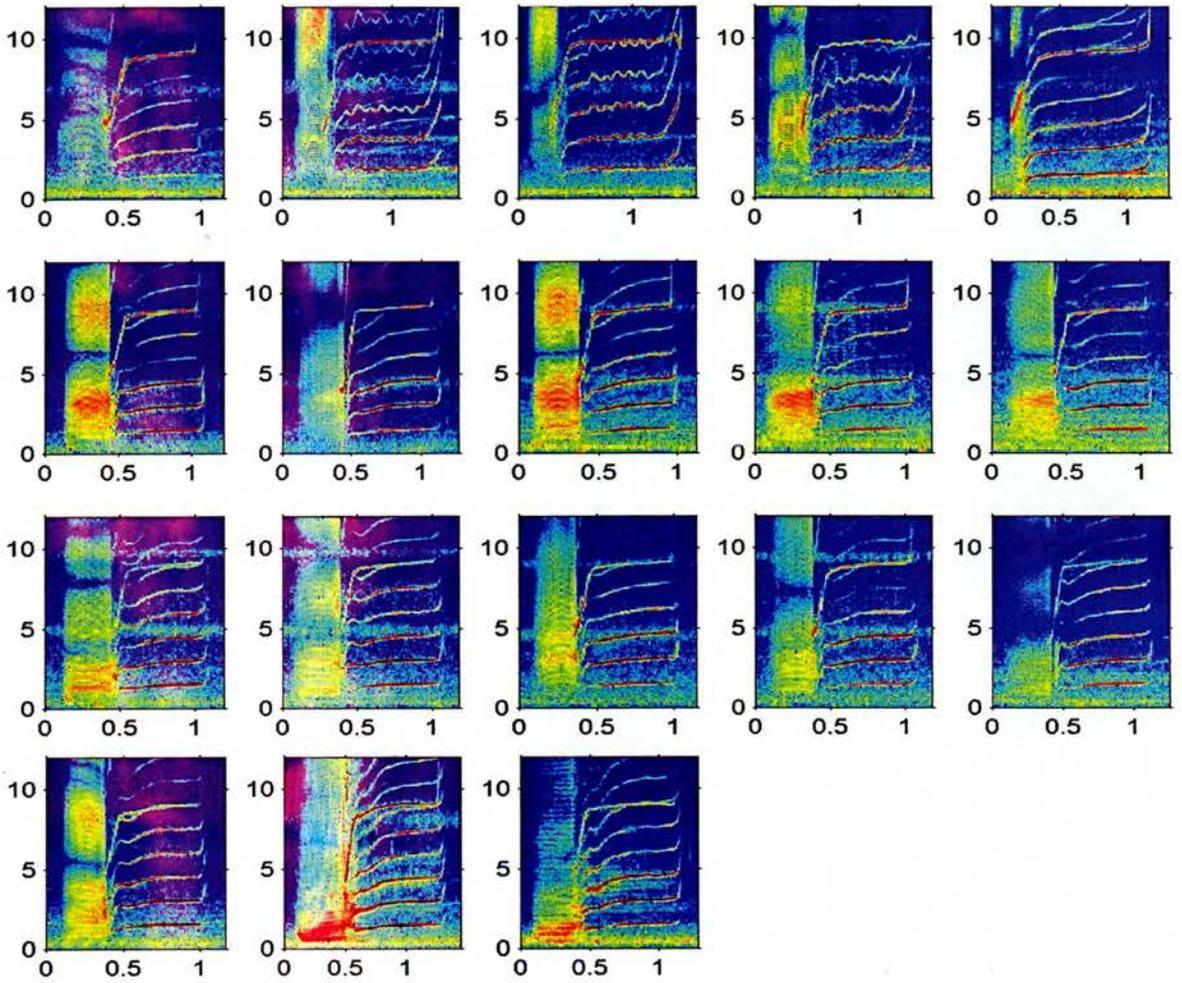
Call Type N4

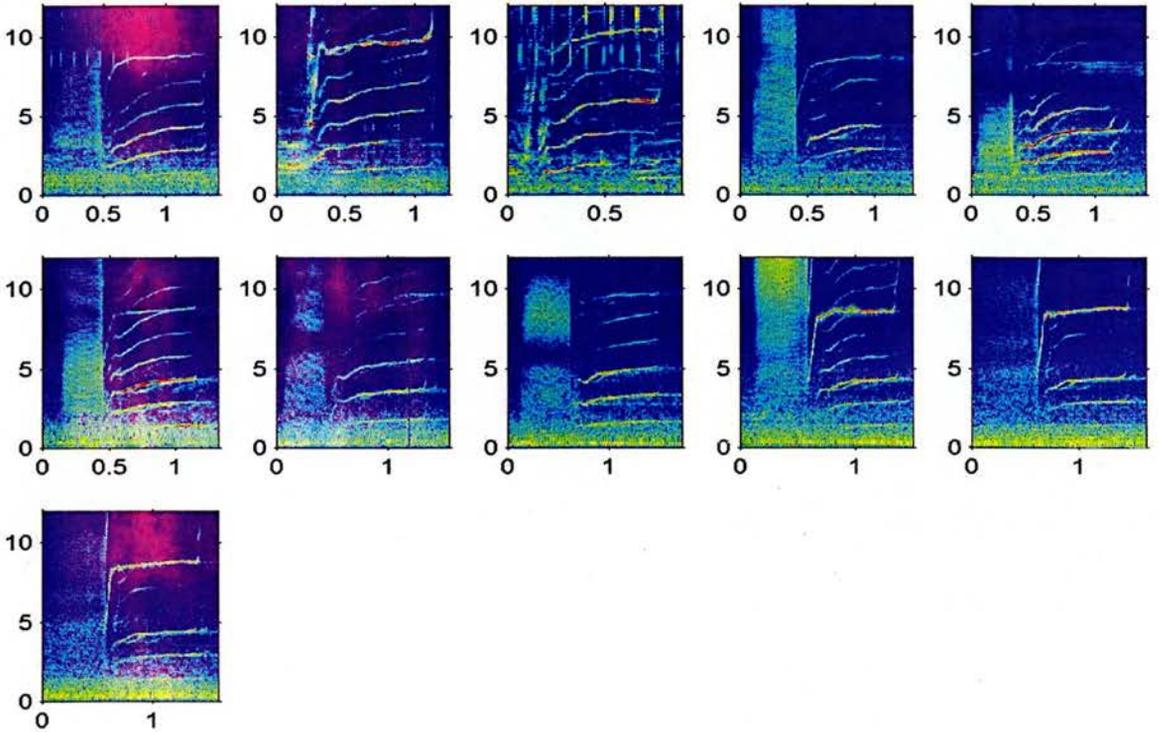
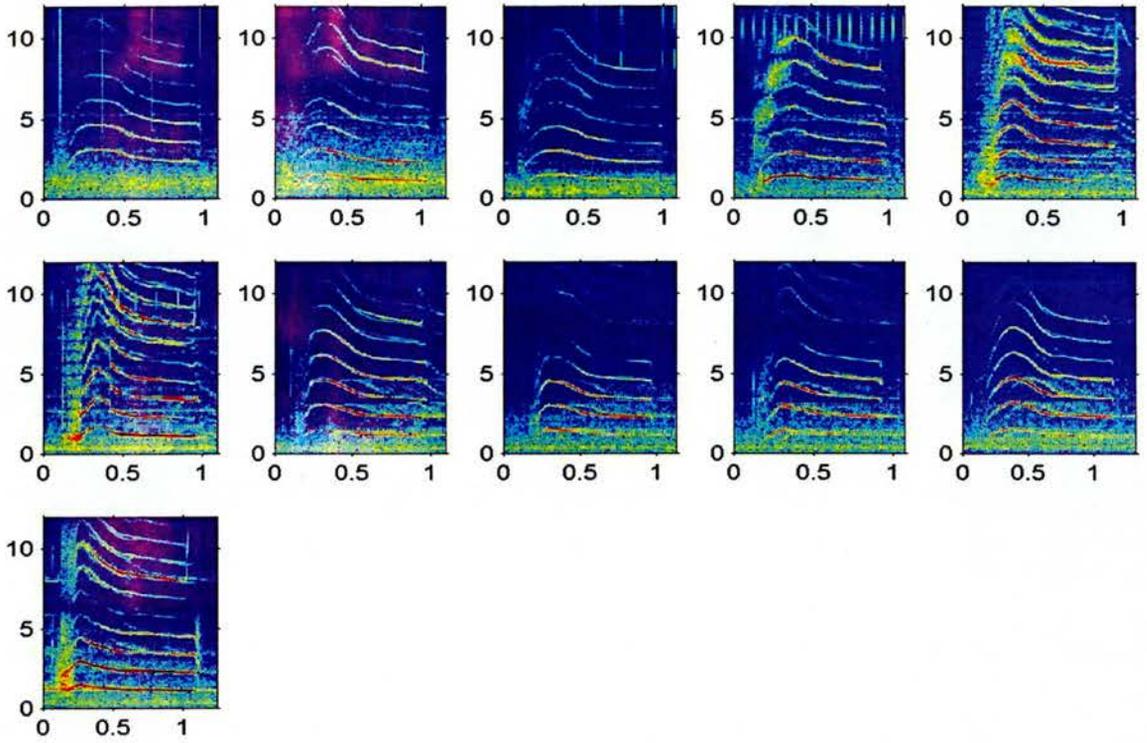


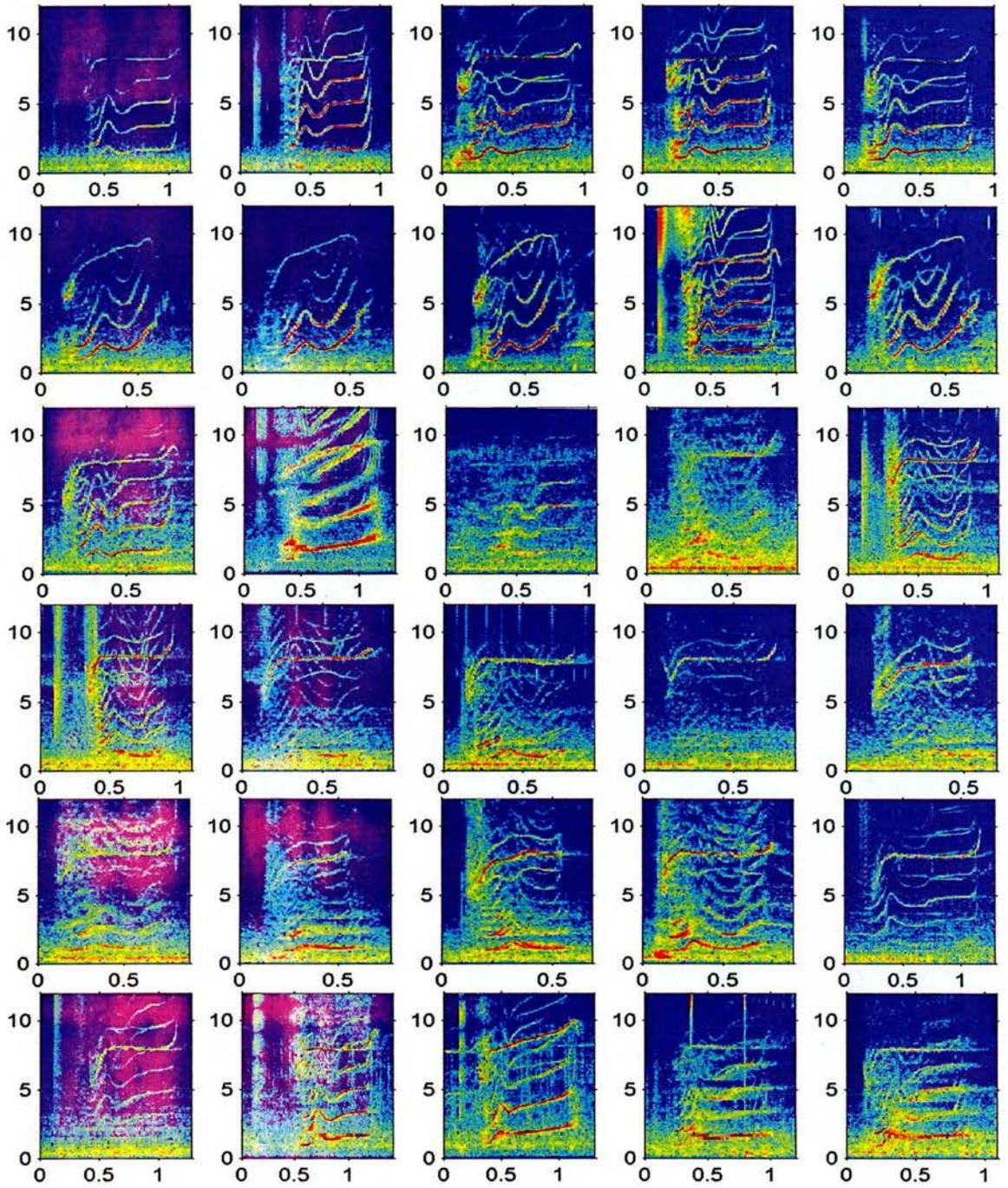
A38 – Matriline A30

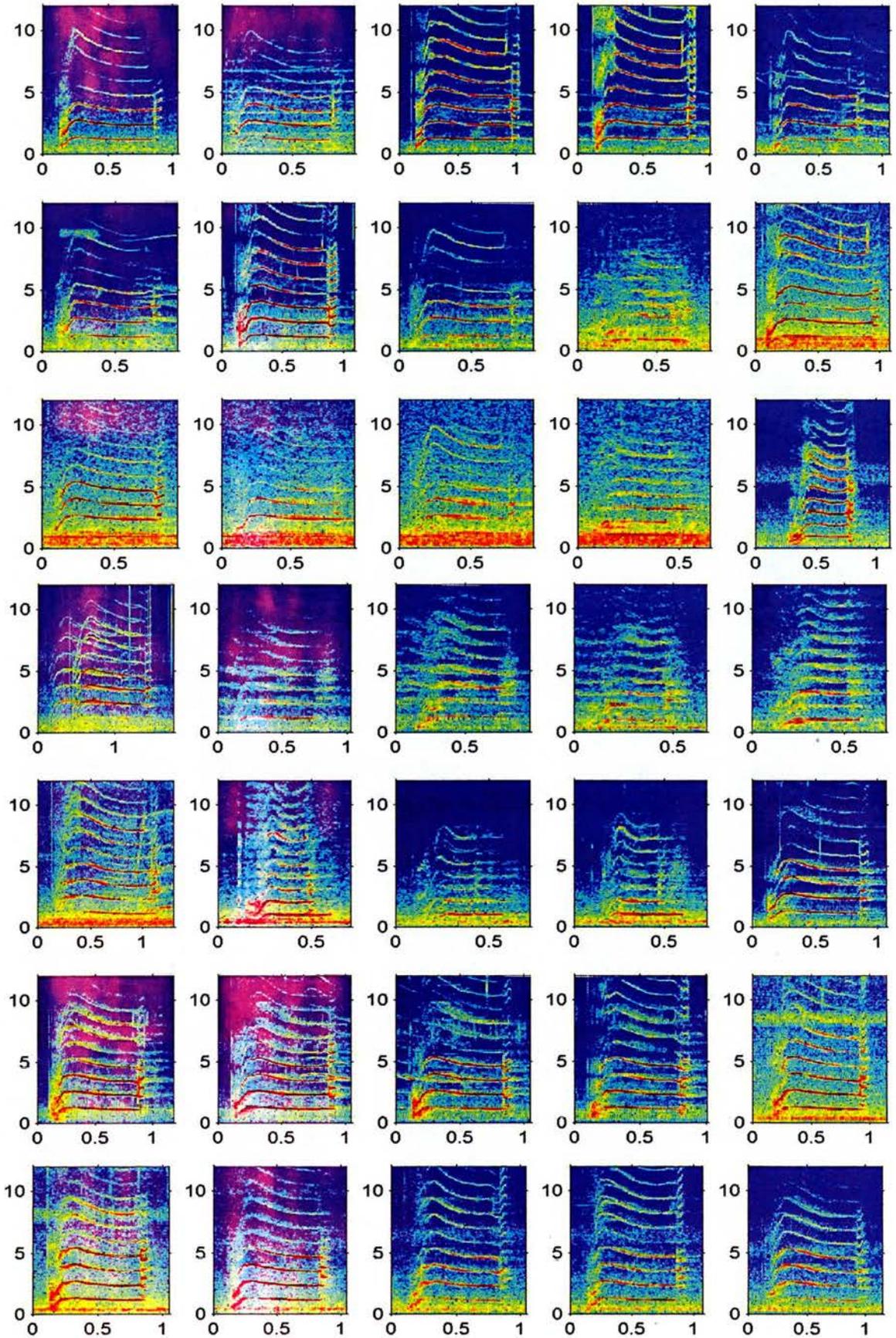
Call Type N4

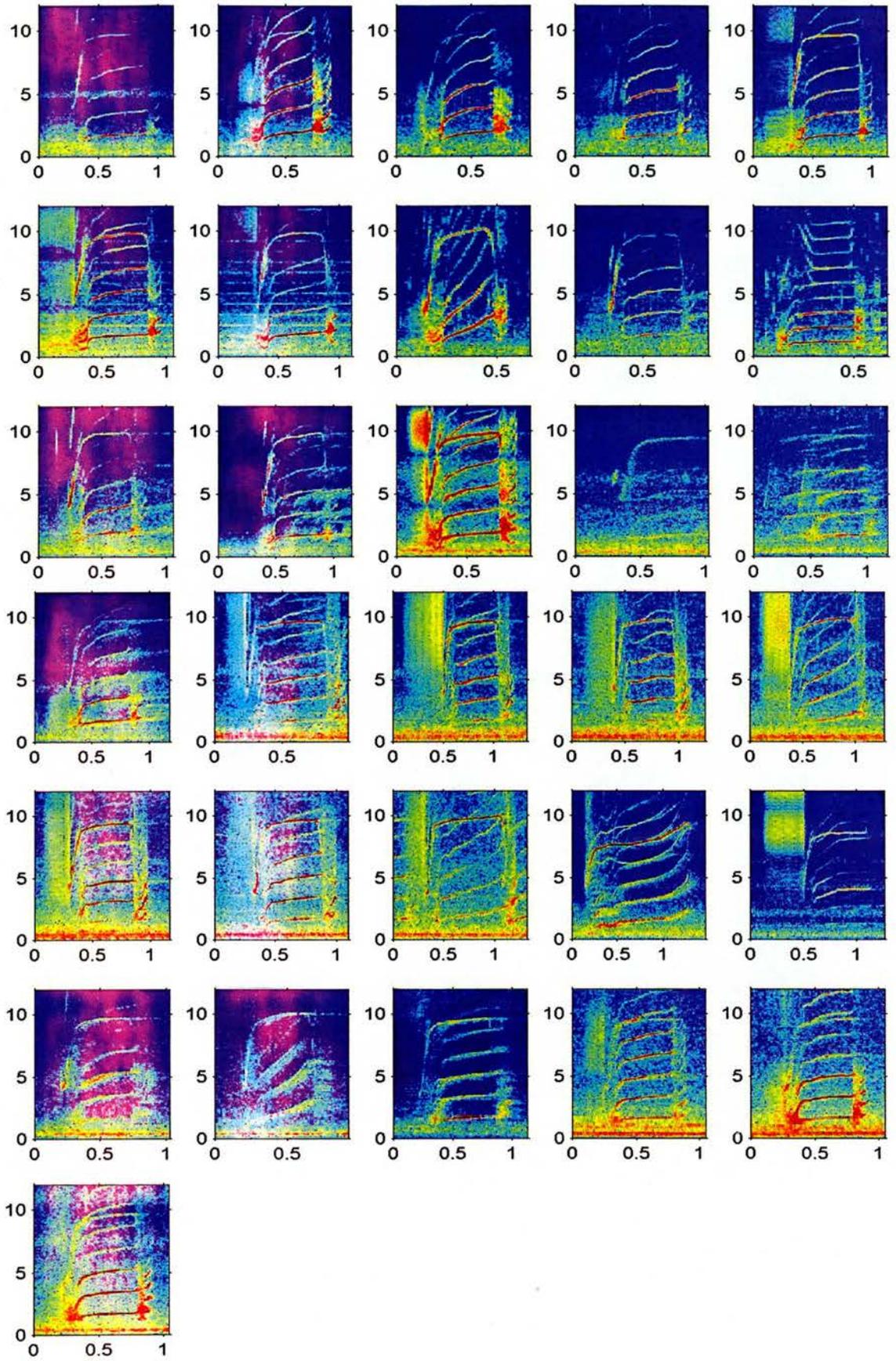






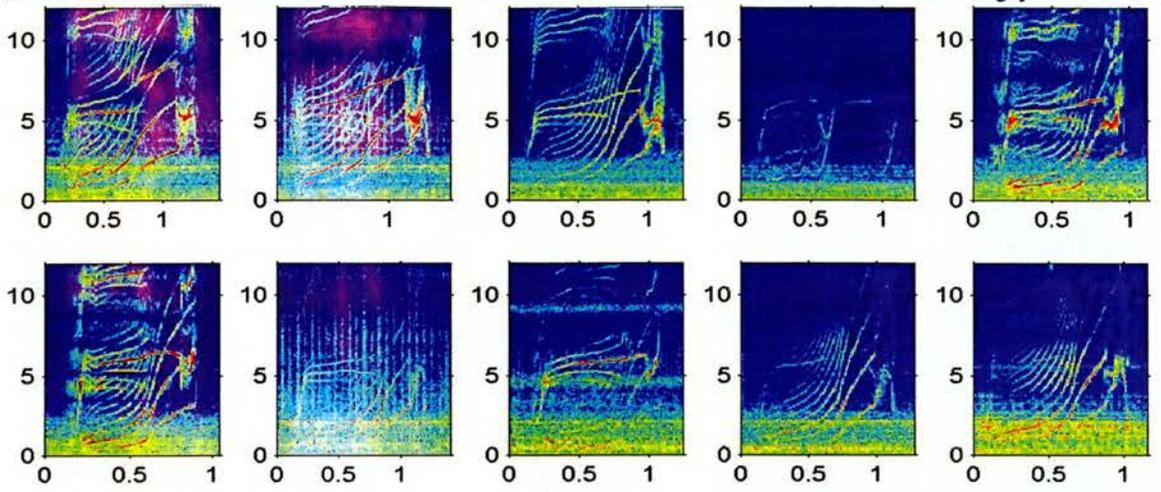






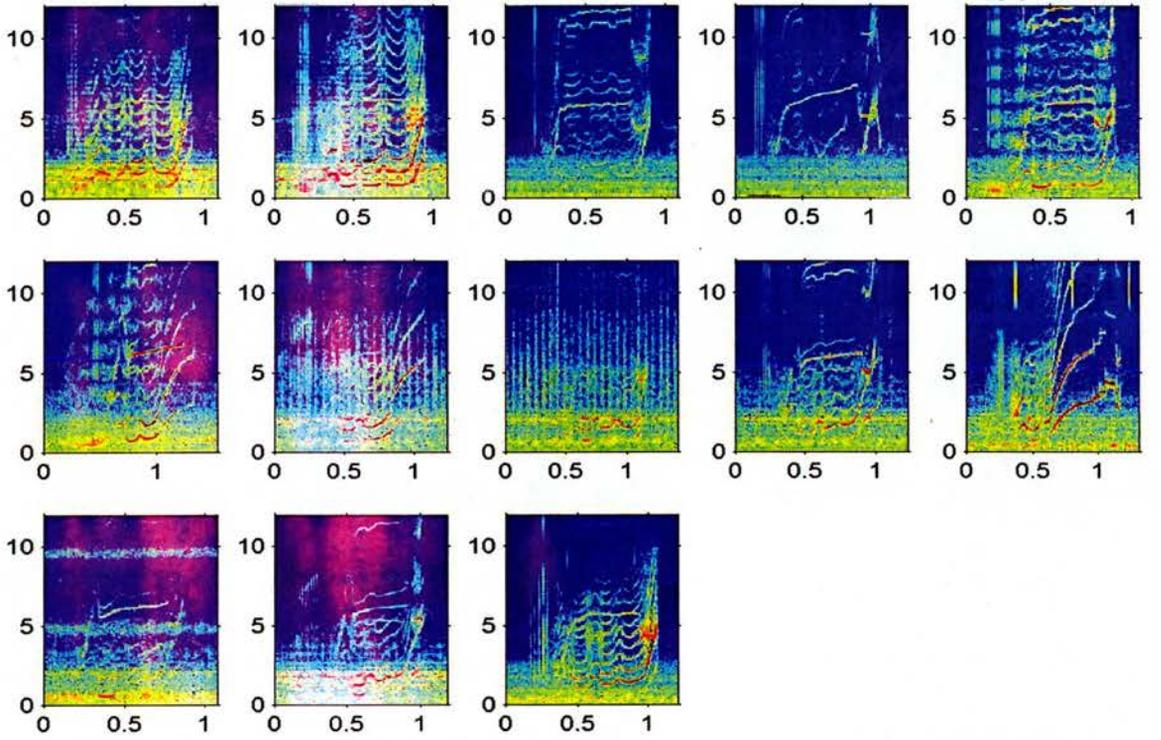
W3 – Matriline W3

Call Type N32



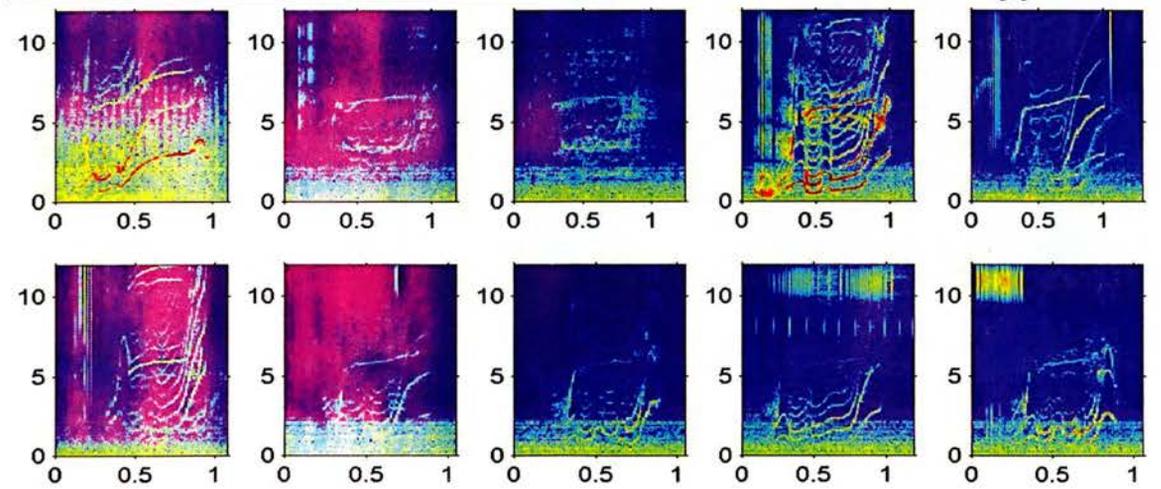
W3 – Continued

Call Type N33



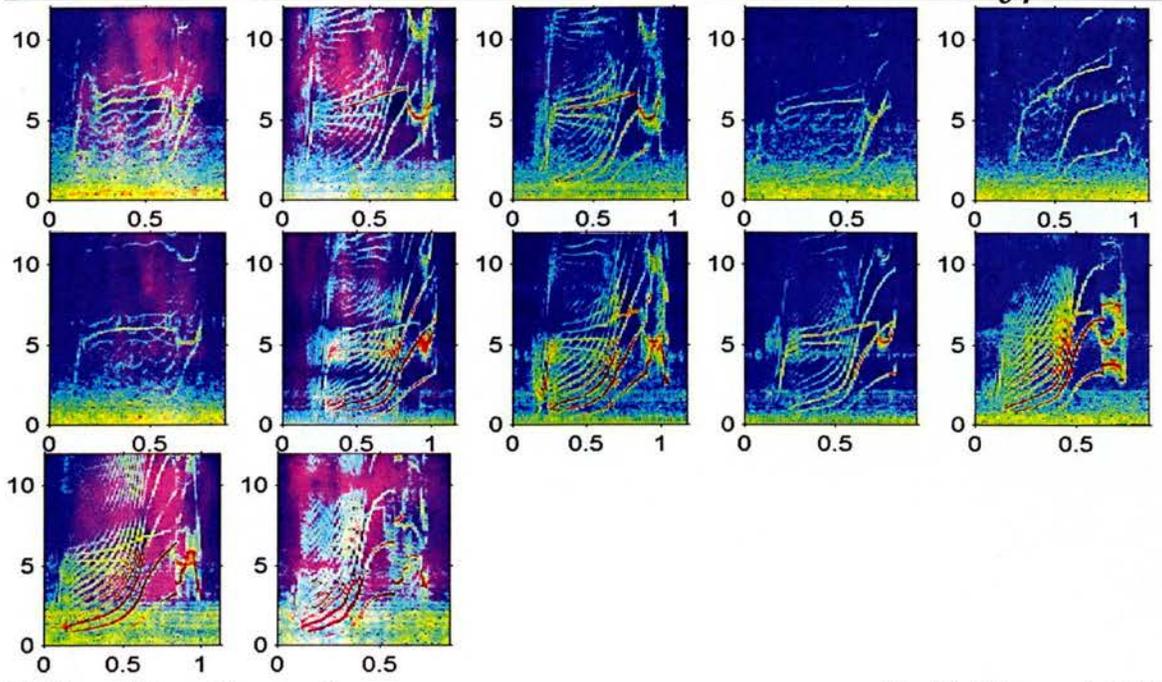
W5 – Matriline W3

Call Type N33



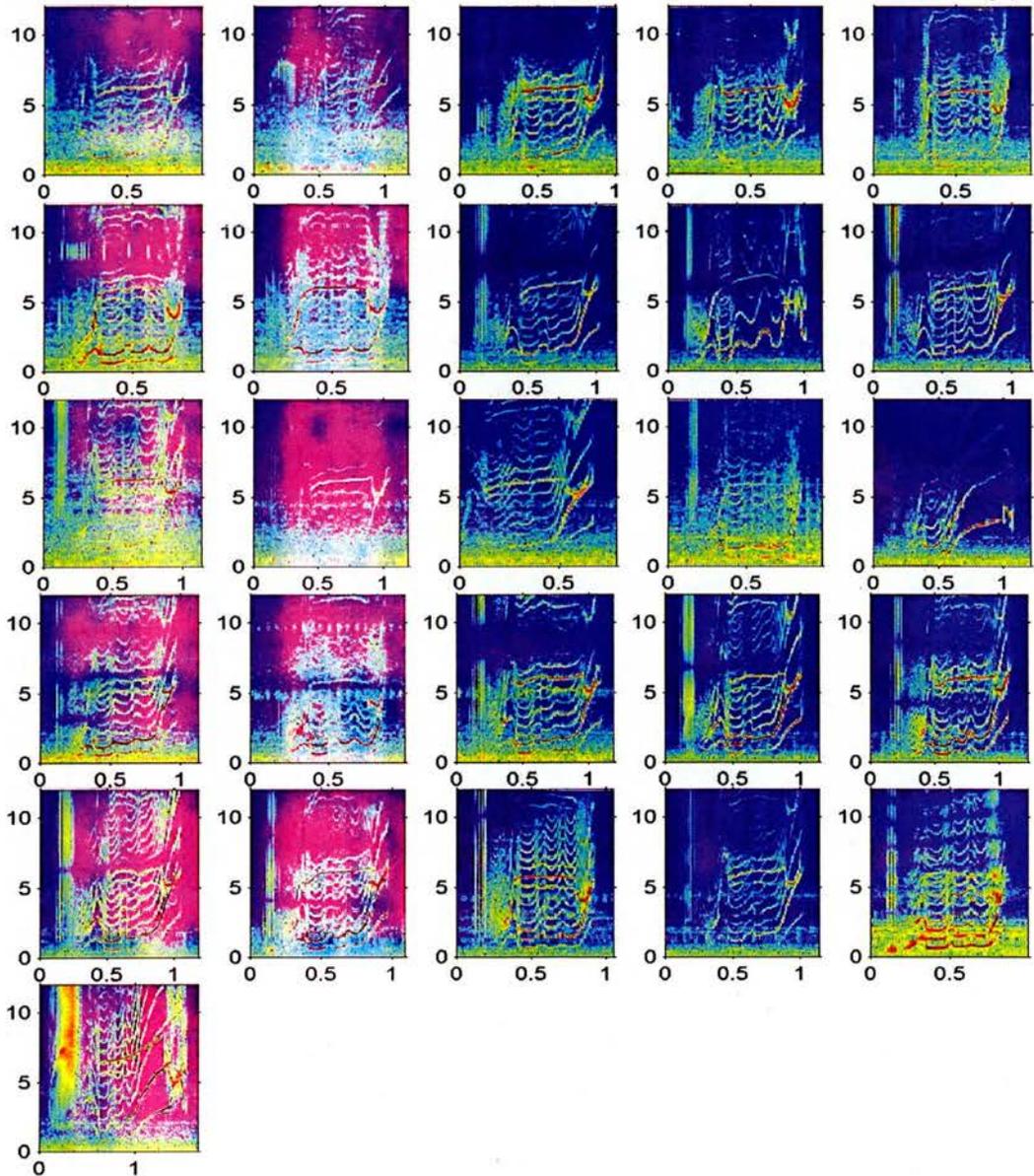
W2 – Matriline W3

Call Type N32



W2 – Continued

Call Type N33



APPENDIX II: SUMMARY OF AVAILABLE FIELD RECORDINGS

This table describes each recording follow with the recording date, focal animal, start time and duration. Chapter 2 contains a description of the methodology used to collect these recordings, and Miller & Tyack (1998) details information about the equipment setup and the development of this recording technique. Focal animals were selected in the field as an individual separated from all other animals present. Follows were terminated when the vessel could no longer be maneuvered to keep that animal in a good recording position as defined in Chapter 2. Follow duration was defined by the duration of the follow where the focal individual was solitary and in a high quality position relative to the array. The number of calls listed in this table describes only those calls ascribed to the focal animal (# Focal Calls) and of those calls, the number rated as being 100% complete (HQ Focal Calls). Follows for which no calls were ascribed to the focal individual are indicated by a star (*) in the follow date column. To describe the ambient noise present for calls included in the analysis, mean RMS (dB re μ Pa) was measured for a silent portion after each high-quality focal vocalization and then averaged for each follow. Any other identified individuals associated with the focal animal during each follow have also been noted here. Finally, any distinct behavioral states of the focal animal during each follow have been included. In general, due to the recording technique most focal animals were traveling during the recording period.

Follow Date	Focal	Start Time (Local)	Duration (min)	# Focal Calls	HQ Focal Calls	Mean RMS (dB re 1 μ Pa)	Other Animals Noted Nearby	Observed Behavioral State
19 August 1998	A8	12:53	38	29	24	118.26	A46, A32, A37	
21 August 1998	A33	13:37	25	1	1	116.46		
25 August 1998	A6	13:25	13	27	26	109.03		
25 August 1998	A38	13:37	36	20	19	111.61		
25 August 1998	A38	14:22	11	37	36	111.29		
26 August 1998	A8	7:57	10	2	2	115.19		
27 August 1998	A12	9:24	46	16	10	113.44		
27 August 1998	A12	10:11	24	17	7	112.59		
28 August 1998*	B8	9:35	42	0	0	N/A		
31 August 1998	A8	10:24	24	9	8	114.61		
31 August 1998	A12	15:14	16	1	1	114.26		
1 September 1998	A33	13:26	32	4	2	107.80		
5 September 1998	A32	11:14	21	5	5	113.47		
5 September 1998	A32	15:17	17	11	11	115.31	A37, A46	
5 September 1998*	A32	15:37	11	0	0	N/A		social, foraging
5 September 1998	A46	15:47	47	4	2	105.56	A32, A37	foraging
5 September 1998	A46	17:07	19	16	16	105.69	A32, A37	
7 September 1998	A46	11:42	25	8	6	110.23	A32, A37	
7 September 1998	A46	13:50	29	9	9	117.13	A32, A36	
7 September 1998	A46	14:36	17	3	3	105.84	A32, A37	
8 September 1998	A46	15:49	8	7	7	117.38	A32, A37	foraging
9 September 1998	A32	11:39	22	43	42	111.17	A37, A46	
9 September 1998	A37	12:27	16	1	1	110.96	A33, A32, A46	
9 September 1998	A37	13:28	13	44	43	118.66	A32, A46	
16 September 1998	A37	13:53	27	3	3	110.76	A46, A32	
16 September 1998	A46	16:18	42	19	11	113.42	A8, A46, A32, A37, A25	
17 September 1998	A8	15:32	39	37	34	112.40	A42, A66	
17 September 1998	A8	16:09	38	13	8	120.36	A42, A66	
17 September 1998	A8	16:42	36	29	17	113.27	A42, A66, A5, A32, A37	
19 September 1998	A46	11:01	40	28	20	116.09	A32, A37	
19 September 1998	A37	11:57	4	3	1	111.39	A32, A46	milling
19 September 1998	A46	12:06	15	11	9	110.22	A32, A37	
19 September 1998	A37	13:37	26	13	13	114.34	A32, A46	
19 September 1998*	A32	16:24	8	0	0	N/A	A37, A46	
19 September 1998	A32	17:04	50	20	20	116.92	A37, A46	milling
20 September 1998	A34	12:32	36	3	3	110.82	A34, A62, A67, A55, A32, A37, A46	
20 September 1998*	A33	13:06	14	0	0	N/A		milling
21 September 1998	A34	8:34	29	9	8	114.68	A55, A32, A37, A46, A67, A62	
21 September 1998	A37	17:06	48	12	11	111.43	A46, A32	milling
23 September 1998	A8	15:10	50	26	23	114.52	A32, A37, A46	
23 September 1998	A8	16:28	17	14	14	117.06	A32, A37, A46	
26 September 1998	A55	14:27	18	22	20	110.69	A34, A62, A67	
26 September 1998	A34	14:53	29	9	9	110.69	A55, A67, A62	
26 September 1998	A62	15:29	4	3	3	112.12	A34, A55, A62, A67	
26 September 1998	A33	15:57	5	1	1	117.12	A34, A55, A62, A68	
22 August 1999	A46	16:21	15	8	2	115.92		
22 August 1999	A46	16:45	20	8	1	115.02	A23, A25, A37	
23 August 1999*	A23	13:41	35	0	0	N/A		
26 August 1999	A46	12:34	25	23	15	115.99		
29 August 1999	A32	16:10	8	8	1	116.53		
12 September 1999	W2	11:56	9	22	11	121.18	W3, W5	
13 September 1999	W5	8:54	4	3	1	120.72	W2, W3	
13 September 1999	W5	9:19	35	16	8	120.66	W2, W3	
13 September 1999	W5	9:48	12	9	0	119.58	W2, W3	
13 September 1999	W2	12:46	38	19	10	119.38	W3, W5	
13 September 1999	W2	17:38	22	5	1	120.87	W3, W5	
14 September 1999	W2	10:27	5	3	2	121.08	W3, W5	
14 September 1999	W2	10:44	24	11	5	120.21	W3, W5	
14 September 1999	W2	16:50	47	35	23	118.95	W3, W5	
14 September 1999	W2	17:42	5	3	0	118.84	W3, W5	
14 September 1999	A27	16:21	17	19	0	120.84	A23, A60, A43, A69	
15 September 1999	A27	9:44	10	6	3	120.22	A23, A60, A43, A69	
15 September 1999	A23	10:44	10	7	6	119.30	A60, A23	
15 September 1999	W5	13:14	12	14	8	118.95	W2, W3	
15 September 1999	W5	14:07	17	2	0	N/A	W2, W3	
15 September 1999	W3	14:32	22	9	5	118.84	W2, W5	
15 September 1999	W3	15:03	17	21	12	118.91	W2, W5	
15 September 1999	W2	16:12	22	4	4	119.14	W3, W5	
15 September 1999	W3	16:47	6	1	0	N/A	W2, W5	
15 September 1999	W3	17:05	19	2	0	N/A	W2, W5	
15 September 1999	W3	16:21	18	13	4	119.23	W2, W5	