VOCAL LEARNING AND DEVELOPMENT IN THE GREY SEAL, HALICHOERUS GRYPUS

Amanda L. Stansbury

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2015

Vocal learning and development in the grey seal, *Halichoerus grypus*
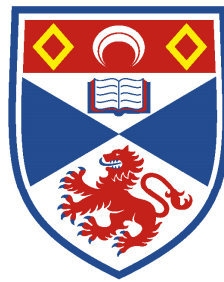
Amanda L. Stansbury

University of
St Andrews

This thesis is submitted in partial fulfilment for the degree of PhD

at the

University of St Andrews

29th May, 2015

1. Candidate's declarations:

I, Amanda Stansbury, hereby certify that this thesis, which is approximately 65,000 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September, 2011, and as a candidate for the degree of PhD in September, 2012; the higher study for which this is a record was carried out in the University of St Andrews between 2011 and 2015.

Date   Signature of candidate

2. Supervisor's declaration:

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

Date   Signature of supervisor

3. Permission for publication:

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 the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

PRINTED COPY

Embargo on all or part of print copy for a period of one year on the following ground(s):

- Publication would preclude future publication

ELECTRONIC COPY

Embargo on all or part of electronic copy for a period of one year on the following ground(s):

- Publication would preclude future publication

Date	Signature of candidate

Date	Signature of supervisor

Abstract

Animal communication systems are complex, but in many species relatively little is known about how they develop. Insight into the development of adult repertoires, considering how factors such as genetics, environment and learning contribute to signal changes, provides a more comprehensive understanding of communication. This thesis documented vocal learning and development in the grey seal. In chapter 3, vocal repertoires of grey seal pups were recorded from birth through their first year, and were compared to calls of other seals across life stages, from pups to adults. By examining call parameters that were similar in animals of the same age, size, and sex, changes that may be attributed to physical development and morphology were identified. Past studies have indicated that seals may also have advanced vocal learning abilities, however the extent of these capabilities was unknown. In this thesis, vocal learning in juvenile grey seals was tested considering both how the seals learn to produce calls as well as gain information from perceived sounds. In chapter 4, grey seals were shown to be capable of vocal usage learning by producing specific calls in specific contexts. In chapter 5, grey seals displayed control over the structure of their calls as they could imitate novel sound models, demonstrating vocal production learning. Lastly, in chapter 6 grey seals gained information from perceived sound signals and used them to their benefit to locate food, demonstrating acoustic comprehension learning. Adult grey seals have complex call repertoires, and in this thesis I have shown how their calls develop with physiological growth and experience, as well as how they can gain information from perceived sound signals.

Most importantly I would like to thank my families for their constant support and love; my parents, Rob and Peggy, Daniel and Felicia, and siblings Katy and Scott, Geoff and Alex. I am especially thankful to Marc, for being my constant companion in every aspect of this journey, despite my insanity over the years.

Contents

**CHAPTER 4: VOCAL USAGE LEARNING IN THE GREY SEAL:
GENERALIZATION OF CALL SIGNAL CLASSES**

List of Tables

List of Figures

Chapter 1

General Introduction

1.1 Introduction

Animal sounds change from birth throughout development. These changes may be attributed to physical maturation and/or learned experience (Sanvito, 2008). Changes due to physical development are predominantly inflexible, falling within limited ranges as they are dependent on morphological restrictions, and may provide information about the signaller such as age, sex and size. Learned signals can have more flexibility and vary between individuals depending on experience.

Janik and Slater (2000) described vocal learning in terms of usage, comprehension and production. Combined, the usage and comprehension of signals are broadly referred to as contextual learning, as they both involve associating an existing signal with a novel context (Janik & Slater, 2000). With usage learning, individuals learn to produce specific sounds in specific contexts as a result of experience. For example, many species have been trained to vocalise when cued (for reviews see Adret, 1992; Janik & Slater, 1997; Schusterman, 2008; Seyfarth & Cheney, 2010). With comprehension learning, individuals learn the contextual significance of sound signals as a result of experience. For example, many species have been trained to respond differentially upon hearing sounds (e.g. Schusterman et al., 1972; Wolski et al., 2003; Gaspard et al., 2012; Stansbury et al., 2014). With production learning, individuals modify the structure of the sounds they produce based on experience with the sounds they perceive, such as with imitation or mimicry.

While many mammalian species have demonstrated vocal contextual learning, evidence for vocal production learning has been accepted for relatively few nonhuman mammals (e.g. Janik & Slater, 1997, 2000; Boughman & Moss, 2003) including: cetaceans (for reviews see Janik & Slater, 1997; Tyack & Sayigh, 1997), elephants (Poole et al., 2005; Stoeger et al., 2012), harbour seals (*Phoca vitulina*) (Ralls et al., 1985) and bats (Esser, 1994). This may be attributed to the strict criteria needed to demonstrate production learning; convincing evidence for production learning relies upon demonstrating that the

animals learned a novel signal that did not previously exist in their repertoire. Using this strict definition, in many cases it is difficult to prove that a sound is truly novel unless it is far outside the typical species repertoire, such as striking instances of mimicking human speech (Ralls et al., 1985; Stoeger et al., 2012). Given such strict criteria for documenting vocal production learning, the paucity of evidence across nonhuman mammals may instead be attributed to the inherent difficulties in documenting vocal learning. Should animals be capable of mimicking such atypical signals, or should imitation of more species specific sounds be more reasonably expected of production learners? Past reviews have argued that production learning may be common to a much wider taxonomic distribution than previously assumed, particularly considering subtler changes in call repertoires (e.g. Janik & Slater, 1997 & 2000; Tyack, 2008). Further examination of mammalian species is needed before conclusions on the prevalence of production learning can be drawn.

Marine mammals are highly adapted for the production and perception of sound. Numerous studies investigating the acoustic capabilities of odontocete species, such as the bottlenose dolphin (*Tursiops truncatus*), indicate they are highly specialized to exploit acoustic information for social communication, prey detection, and navigation (e.g. Au, 1980; Janik & Slater, 1998; Janik et al., 2006). Comparatively fewer investigations have considered the abilities of other marine mammals, such as pinnipeds, which also appear to have advanced vocal capabilities. Studying these other marine mammal species is of interest as they produce sound using the larynx similarly to most other mammals (including humans), while odontocetes generate sound using phonic lips through a system of air sacs and fat compartments (Cranford, 2000). Examining the vocal abilities of pinnipeds provides a better comparative understanding of the capabilities of mammals.

Usage learning, in which vocalisations have been placed under stimulus control, has been demonstrated in California sea lions (*Zalophus californianus*) (Schusterman & Feinstein, 1965; Schusterman & Balliet, 1970; Schusterman, 1978), Pacific walruses (*Odobenus rosmarus divergens)* (Schusterman & Reichmuth, 2008), harbour seals (Schusterman, 2008)*,* and grey seals

(*Halichoerus grypus*) (Shapiro et al., 2004). Usage learning has also been shown to occur in wild populations of Southern elephant seals (*Mirounga leonina*) (Sanvito et al., 2007). Harbour seals and Pacific walruses are additionally capable of voluntary control over vocal plasticity (Schusterman & Reichmuth, 2008; Schusterman, 2008). This indicates they have flexibility and control over call production. Production learning has also been anecdotally observed in a harbour seal which spontaneously mimicked human speech sounds (Ralls et al., 1985). Given these past studies, further investigation into vocal development and learning, particularly production learning, of pinnipeds is warranted.

Here I will briefly discuss the natural history and acoustic behaviour of pinnipeds. Grey seals (*Halichoerus grypus*) are amongst the most vocal pinniped species, and their vocal repertoire will be highlighted.

1.2 Brief Introduction to Pinnipeds
1.2.1 Pinniped Natural History

There are 33 living pinniped species from three families; Phocidae, the 'true seals', Otariidae, the 'eared seals' consisting of sea lions and fur seals, and Odobenidae, the walrus (Berta & Churchill, 2012). All are semi-aquatic and predominantly distributed throughout the colder waters of the Northern and Southern hemisphere. Pinnipeds spend the majority of their time in water, and variable amounts of time hauled out terrestrially. They move to land for extended periods to give birth, molt and mate.

Pups are usually born on land, nursing from their mothers for a time before weaning. Maternal care is variable across pinniped species; some nurse pups for as little as four days before weaning the pups, while others have maternal care for up to three years (Bowen, 1991). During the nursing period, pups typically gain a large amount of weight, and after weaning go through a fasting period before beginning to forage independently.

Pinnipeds have variable social structures; when at sea, individuals are thought to typically remain isolated with only temporary associations, but terrestrially they may congregate in large, dense populations (Pomeroy et al.,

1994; Pomeroy et al., 2000). There is some question as to how isolated individuals are at sea; some evidence suggests that individuals may regularly meet (Lidgard et al., 2012). However, this result is questionable as the animals were tracked with acoustic tags which were perceptible to the seals and could have inadvertently attracted individuals (this is further discussed in chapter 6). Terrestrially, some species return to the same haul outs regularly; for example, in grey seals many females return to the same pupping site across seasons (Pomeroy et al., 1994; Pomeroy et al., 2000).

1.2.2 Pinniped Vocal Behaviour

All three pinniped families vocalise both in-air and underwater (Schusterman & Van Parijs, 2003). It is hypothesized that seals may passively use sound for orientation or navigation (Schusterman et al., 2000; Evans et al., 2004), however the only known role of vocal behaviour is for social communication (Schusterman, 2008). Calls have been found to play a role in individual identification, mother-pup reunions, agonistic and/or territorial interactions, and mate attraction.

Several species of pinniped have vocalisations that vary significantly between individuals, and could potentially be used to gain information about caller identity (Table 1.1). Although call parameters differ between individuals in these species, they are not necessarily used by the animals for recognition. However, some pinniped species have been shown to use call features for recognition with playback studies, such as in harbour seals (Renouf, 1985), Subantarctic fur seals (*Arctocephalus tropicalis)* (Roux & Jouventin, 1987; Charrier et al., 2002), Australian fur seal (*Arctocephalus pusillus doriferus*) (Tripovich et al., 2008), Australian sea lions (*Neophoca cinerea)* (Charrier et al., 2009), Northern fur seal (*Callorhinus ursinus*) (Insley, 2000), California sea lions (Schusterman et al., 1992; Hanggi, 1992), and Galapagos sea lions (*Zalophus wollebaeki*) (Trillmich, 1981).

The ability to identify individuals based on vocalisations, in addition to using other cues such as vision and olfaction, is hypothesized to facilitate mother pup reunions. In several species, pups are heavily dependent on

Table 1.1. List of pinniped species with significant inter-individual differences in call parameters.

Common Name	Latin Name	Source
Grey Seals	*Halichoerus grypus*	Caudron et al., 1998; McCulloch et al., 1999
Harp Seals	*Phoca groenlandica*	Van Opzeeland & Van Parijs, 2004
Harbour seals	*Phoca vitulina*	Renouf, 1984; Hanggi & Schusterman, 1994; Khan et al., 2006
Weddell seals	*Leptonychotes weddellii*	Terhune & Dell Apa, 2006; Collins et al., 2006
Leopard seals	*Hydruga leptonyx*	Rogers & Cato, 2002
Northern elephant seals	*Mirounga angustirostris*	Shipley et al., 1981
California sea lions	*Zalophus californianus*	Gisiner & Schusterman, 1991
South American sea lions	*Otaria flavescens*	Fernández-Juricic et al., 1999
Galapagos sea lions	*Zalophus wollebaeki*	Trillmich, 1981
Australian fur seals	*Arctocephalus pusillus doriferus*	Stirling & Warneke, 1971; Tripovich et al., 2005; Charrier & Harcourt, 2006
Subantarctic fur seals	*Arctocephalus tropicalis*	Roux & Jouventin, 1987; Page et al., 2002
Antarctic fur seals	*Arctocephalus gazelle*	Page et al., 2002
New Zealand fur seals	*Arctocephalus forsteri*	Stirling & Warneke, 1971; Page et al., 2002
Northern fur seal	*Callorhinus ursinus*	Insley, 1992
Galapagos fur seal	*Arctocephalus galapagoensis*	Trillmich, 1981
Atlantic walrus	*Odobenus rosmarus rosmarus*	Stirling et al., 1987
Pacific walrus	*Odobenus rosmarus divergens*	Kastelein et al., 1995

maternal care during a brief and rapid maturation period after birth, leading to a high energetic cost of nursing on the part of the mother (Costa et al., 1986; Bowen, 1991). Pupping can occur in dense colonies, with mothers leaving their offspring periodically to return to sea. While this foraging strategy most frequently occurs in otariids, it has also been observed in some phocid species (Boness & Bowen, 1996). Individualized vocalisations could aid females in

identifying and returning to their pups. In some species, mothers respond preferentially to playbacks of their offspring rather than neighbouring infants or strangers, indicating vocal cues could be used to identify pups (e.g. Insley, 2000; McCulloch et al., 1999). Although pup calls may change with age, in some species mothers are able to recognize successive versions of their pup's call as they develop (e.g. Charrier et al., 2003). Recognition is not limited to mothers identifying pups; Galapagos sea lion (Trillmich, 1981) and California sea lion (Schusterman et al., 1992) pups respond preferentially to their mothers' vocalisations. Age can affect the ability to identify vocalisations; in the Australian sea lion pup recognition of the mother's vocalisation is delayed for two months (Pitcher et al., 2009). Vocal recognition is also not limited to the time period when pups are dependent on their mothers; Northern fur seals respond differentially to their offspring's vocalisations into adulthood (Insley, 2000).

Not all pinniped species use vocal recognition in mother pup identification. In Weddell (*Leptonychotes weddellii*) (Van Opzeeland et al., 2011) and one colony of grey seals (McCulloch et al., 1999), females respond equally to playbacks of their own and other pups. These species may use other cues such as location and/or olfaction when returning to pups. Errors occur as several cases of adoption and allo-suckling have been recorded (see Bowen, 1991 for review). While misidentification may indicate errors in recognition, it is also possible there are adaptive benefits to nursing other pups such as kin selection or milk evacuation (Roulin, 2003).

The ability to discriminate between individuals also influences agonistic and territorial interactions. When individuals can be identified, unnecessary conflict can be averted by weighing the costs and benefits of aggressing against a known conspecific. For example, Weddell seals live in a fast-ice environment where territorial males defend established breathing holes. One male-specific vocalisation, a loud, long, descending frequency call referred to as a 'trill', has been observed during dives from territorial males defending breathing holes (Kooyman & Kooyman, 1981; Thomas & Stirling, 1983; Thomas et al., 1983; Bartsh et al., 1992; Oetelaar et al., 2003). The calls are individually unique

(Thomas & Stirling, 1983), allowing males to identify each other while defending territory (Terhune & Dell Apa, 2006; Terhune et al., 2008). Similarly, male Northern elephant seals (*Mirounga angustirostris*) respond differentially to calls of neighbouring males depending on previous territorial interactions (Casey et al., 2013). Responses did not vary by rival male size (Charrier et al., 2013), suggesting that morphological cues did not mediate territorial interactions. Instead, male calls appeared to indicate identity, and individuals remembered previous social interactions to evaluate the threats before aggressive interactions occurred. This facilitates the assessment of costs and benefits to potential interactions and prevents unnecessary conflict.

Aggressive calls may not always be associated with specific individuals. Rogers (2003) reviewed the literature concerning vocalisations of male phocid seals and several species produce specific call types in agonistic interactions, usually in male-male conflicts. For example, when wild harbour seals were exposed to an aggressive call type, the 'roar', of conspecifics they showed increased aggressive behaviours, such as approaching the speaker and flipper slapping (Hayes et al., 2004). This aggressive behaviour occurred more frequently with high amplitude 'roars', suggesting that call intensity may be used as a graded signal in agonistic interactions (Hayes et al., 2004). While these signals are not necessarily associated with specific individuals, this does not mean the calls could not additionally provide information about caller identity.

Male vocalisations also play a role in mate attraction (Rogers, 2003). In phocid seals, the acoustic repertoires of males during the mating season are more varied and complex than at other times (Green & Burton, 1988a & b; Morrice et al., 1994; Van Parijs et al., 1997). For example, male harbour seals restrict their travel range and increase diving and acoustic behaviour when females are in oestrus, presumably as a display to attract mates (Van Parijs et al., 1997). These calls are emitted in varied, recognizable sequences which fit Thorpe's (1964) definition of bird song. Consequently, in several seal species male mating vocalisation displays have been referred to as song (Thorpe, 1964; Stirling, 1973; Ralls et al., 1985; Green & Burton, 1988b; Morrice et al., 1994).

1.2.3 Grey Seal Vocal Repertoire

Adult grey seals produce a variety of aerial and underwater vocalisations that have been categorized and labelled by several names including (but not limited to): 'kataro', 'gurgle', 'moan', 'growl', 'yodel', 'rup', 'rupe', 'trrot', 'buzz', 'click', 'moo', 'hiss', 'knock', 'hoot', and 'roar' (Hewer & Backhouse, 1960; Schevill et al., 1963; Schusterman et al., 1970; Oliver, 1978; Asselin et al., 1993; McCulloch, 1999; Shapiro et al., 2004). While the vocal abilities of grey seals have been reported in several studies, only two have conducted a thorough investigation into the full extent of their vocal repertoire. In a doctoral thesis, McCulloch (1999) recorded aerial and underwater vocalisations from two pupping beaches during three breeding seasons in Scotland (Isle of May) and Nova Scotia (Sable Island), and underwater calls from another site in Scotland (Abertay Sands). Calls were initially categorized subjectively based on visual evaluation of the spectrogram and/or auditory characteristics, and then given to human observers to further group. Calls were classified with high inter-observer reliability into six aerial (labelled A through F) and ten underwater (labelled 1 through 10) types. In addition to numeric or alphabetic categorization, descriptive names were applied for some call types (McCulloch, 1999). Asselin et al. (1993) recorded thirty-six hours of underwater vocalisations from a haul-out and breeding site in Nova Scotia (Gulf of St. Lawrence, Amet Island), and categorized calls using the same method. Six call types were identified and categorized using descriptive names (Asselin et al., 1993).

Most calls were classified and described similarly across studies such as the 'rup' (call type 1) (Asselin et al., 1993; McCulloch, 1999), 'rupe' (call type 5) (Asselin et al., 1993; McCulloch, 1999), 'trrot' (call type 3) (Asselin et al., 1993; McCulloch, 1999), 'growl' (call type 9) (Asselin et al., 1993; McCulloch, 1999; Shapiro et al., 2004) and 'click' (Schevill et al., 1963; Schusterman et al., 1970; Oliver, 1978; Asselin et al., 1993). However, there were several differences between studies which can most likely be attributed to variable categorization and labelling. Across studies some calls appear to be described similarly, but categorized by a different name. For example, the call labelled as a 'moan' (call type 7) by McCulloch (1999), Shapiro et al. (2004) and Schusterman (1970), a

'hoot' by Hewer and Backhouse (1960), and a 'roar' by Asselin et al. (1993) are described similarly (all long duration, frequency modulated periodic calls with harmonics). When comparing between studies perhaps these calls should be classified as the same type.

McCulloch (1999) appeared to split call types where Asselin et al. (1993) consolidated them. For example, McCulloch (1999) identified call type 2, described as a single element call with wide frequency range, brief duration, produced in multiple-call sequences. This is very similar to her call type 1, the 'rup', the difference appearing to be duration and time between elements. Asselin et al. (1993) noted considerable variability in duration and time between 'rups', but still classified them as the same call type. While McCulloch (1999) appeared to split the calls into different categories due to this variability, Asselin et al. (1993) appeared to consolidate them into a single call type.

While most calls appear to be single element calls, many are repeated in sequences (defined as calls produced without a break of more than two seconds, Asselin et al., 1993) such as 'rups', 'knocks' (call type 8), 'trrots' and 'clicks' (Asselin et al., 1993; McCulloch, 1999). Notably, the 'rupe' (call type 5) is repetitive, multiple element call (Asselin et al., 1993; McCulloch, 1999). One element of the call consists of a noisy pulse which appears to be very similar to the 'rup'. While McCulloch (1999) did not describe the difference between the 'rup' and first element of the 'rupe', Asselin et al. (1993) noted more harmonic structure in the element of the 'rupe'. However, examples of 'rups' and 'rupes' provided by McCulloch (1999) appear to be very similar and were not differentiated by harmonic structure, suggesting that in her analysis perhaps one element of the 'rupe' may be described as a 'rup'. This is further discussed, and examples are provided, in chapter 3. The second element of the call consists of a more periodic, harmonic component that is typically longer in duration. The order of these two elements appears to be flexible, and varies in several parameters such as element duration, intercall duration, frequency structure (such as an upsweep or downsweep), and sequence length (Asselin et al., 1993).

Only one type of vocalisation has been reported for the grey seal pup (Asselin et al., 1993; Caudron et al., 1998; McCulloch et al., 1999). Although described using many names ('pup call', 'pup begging call', 'mother attraction call', 'bleat') in all cases the vocalisation is frequency modulated at the beginning and end of the call, and usually contains a flat contour at the centre of the call (McCulloch et al., 1999). It has a harmonic structure (up to eighteen) with the highest amplitude in the fundamental, first, or second harmonic (Asselin et al., 1993).

While the vocal repertoire of grey seals has been documented in these studies, very little is known concerning the biological function of the different call types. As previously discussed, observations and playback experiments have shown that in general seal calls are used for social communication. Grey seals are also most vocal during the pupping and breeding seasons, with some call types occurring predominantly (such as McCulloch's call types 1 and 8) during the mating period, suggesting they may play a role in mate selection (Asselin et al., 1993; McCulloch, 1999). However, the only known role of grey seal calls is for mother-pup recognition (McCulloch et al., 1999; McCulloch & Boness, 2000). When played either their own or other seal pup calls, grey seal mothers in one colony (Sable Island, Nova Scotia, Canada) responded with more head turns and body movements to her own pup's call. However, this did not occur with mothers on another seal colony (Isle of May, Firth of Forth, Scotland) (McCulloch et al., 1999; McCulloch & Boness, 2000).

As part of a PhD thesis, another investigation has also played back grey seal calls. Götz (2008) played captive grey seals the underwater call types previously recorded by McCulloch (1999) including: 'moans', 'rupes', 'rups', 'growls', 'type 10 calls', and 'knocks', during a feeding experiment. While the seals were not deterred from feeding by any of the conspecific call types, the seals were more likely to approach the speaker when played 'moans', 'rupes' and 'rups' (Götz, 2008). However, it is unknown as to why the seals approached for these call types. These call types could have attracted the seals for a variety of reasons, ranging from aggression to affiliation. While pup calls of grey seals

allow for individual identification and facilitate mother-pup recognition, the biological function of the different call types remains unclear.

1.3 Thesis overview

Animal communication systems are complex and it is difficult to evaluate how adult repertoires arise through physical (including neurological) development and learning. In chapter 2, I review the current literature concerning call development and learning in mammals. Throughout the rest of this thesis, I examine vocal development and learning in the grey seal. Previous research concerning vocalisations of pinnipeds has focused on documenting vocal repertoires and the behavioural context of calls (for a review see Rogers, 2003). However, little information is available on call development. In chapter 3, the ontogeny of grey seal vocalisations will be examined, documenting the repertoires of pups from birth through their first year and comparing calls across life stages, from pups to adults. Past research investigating call learning found that grey seals are capable of vocal usage learning by matching call types (Shapiro et al., 2004). However, the seals could only correctly categorize known stimuli, and were unable to generalize the ability to classify novel examplars of sound types. In chapter 4, vocal usage learning will be examined in a juvenile grey seal which was successfully trained to discriminate novel calls and classify them into call types. Previous anecdotal evidence indicated that seals were capable of imitating novel sounds. In chapter 5, vocal production learning will be investigated in grey seals trained to imitate novel sound sequences. Lastly, some evidence suggests that pinnipeds are capable of associating acoustic signals with food (e.g. Jefferson & Curry, 1996). In chapter 6, acoustic comprehension learning will be investigated in grey seals using an anthropogenic sound signal, a fish pinger, for prey detection in a search task to find hidden food.

In summary, the aim of this thesis is to document the vocal skills of grey seal pups and how their calls develop and change over time. By examining call changes that occur across seals, trends that occur in animals of the same age, size, and sex may be attributed to physical growth, through changes in

morphology and neural development. Changes that occur within individual animals, such as when calls are produced, call structure, and how perceived sounds are used can be attributed to learning. Considering both physical development and learning will result in a more comprehensive understanding of how the complex vocal repertoires of pinnipeds arise.

Chapter 2
Vocal Flexibility and Development in Mammalian Communication:
A Review

2.1 Introduction

Animal communication systems are complex, however relatively little is known about their development. Understanding how adult repertoires develop, considering changes occurring from learning, physical and neural development, provides a more comprehensive understanding of adult communication. Previous studies investigating vocal development in non-human animals have been predominantly based on birds, particularly the taxa capable of vocal production learning (the oscine passerines (songbirds), psittaciformes (parrots) and trochiliformes (hummingbirds)). Here I examine the current literature on vocal development concerning mammals, with an emphasis on pinnipeds. Although the term 'vocal' implies that signals are created from the manipulation of vocal cords, in this review the term is used broadly to describe any biological sound produced by the vibration of body tissue. First, I discuss the call changes that can most likely be attributed to learning, and then those that can most likely be attributed to physical growth, through changes in morphology and neural development.

2.2   Vocal Learning in Mammals
2.2.1   Introduction

Vocal learning is the process by which individuals modify the signals they produce and their responses to perceived signals based on external auditory information and feedback. As previously discussed, this can affect contextual learning through the usage (individuals learn to produce signals in a new context from experience), and comprehension (individuals learn new contextual significance, or meaning, of any sound signal from experience) of signals, as well as the production (individuals learn to modify signal structure from experience with other sounds) of calls (Janik & Slater, 2000). Here, I will discuss the current literature concerning vocal learning in mammals. First, contextual learning will be discussed, separating usage and comprehension

learning. For vocal usage learning, the ability to learn specific contexts to produce specific calls is common among birds and mammals. As it has been widely documented, vocal usage learning will only be briefly discussed and will highlight studies similar to my own research; those working with trained, captive animals, particularly pinnipeds.

Next, vocal comprehension learning will be discussed. Traditionally, comprehension learning refers to information received from con- and hetero-specific vocalisations (Janik & Slater, 2000). However, individuals can similarly gain information from any sound signal in their environment. Rather than vocal comprehension learning, this is described as acoustic comprehension learning as it encompasses learning contextual significance from any perceived sound signal. In this section I briefly discuss studies examining vocal comprehension learning of biological sounds, both from conspecifics and heterospecifics, as well as acoustic comprehension learning of anthropogenic sounds. Most research concerning the impact of anthropogenic signals has been limited to detrimental effects, however there are several ways animals could also benefit from acoustic signals through comprehension learning.

Lastly, vocal production learning will be reviewed. Production learning has been well demonstrated in birds and humans, but is accepted in very few nonhuman mammals. Recently there has been considerable interest in production learning, partly due to the discrepancy between the advanced language capabilities of humans in comparison to other mammals, especially primates. The subject has been well reviewed (e.g. Janik & Slater, 1997; Schusterman, 2008; Tyack, 2008; Seyfarth & Cheney, 2010), however our understanding of production learning is still relatively limited. This is partly attributed to the difficulties of demonstrating production learning when call changes can be explained by other variables such as genetics, maturation, motivation, improvisation, and environmental changes. Additionally, all of the accepted vocal learning species have learned a novel signal that falls outside species-typical repertoire. Rather than using the arguably advanced ability to mimic species-atypical sounds as a standard for measuring production learning, perhaps the capabilities of some species may be better estimated by looking at

subtler changes such, as with convergence or divergence. Additionally, most current research looks at production learning as a black or white capability; it is either present or absent in a species. However, animal communication systems are complex and vocal production learning abilities may occur on a continuum. The currently accepted vocal learning species have all displayed arguably advanced control over vocal production by varying the frequency of their calls by manipulation of the sound's source (the larynx in most mammals) and filter (the vocal tract in most mammals). However, other species may have control over a less complicated aspect of the vocal tract, the respiratory system. Current 'non-learning' species may be able to manipulate call aspects such as amplitude, duration, and call rate which have not been considered in most of the vocal production literature (Janik & Slater, 1997 & 2000). Here, I will review the current literature on vocal production learning in depth and discuss how the ability might be measured to reflect varying capabilities across species.

2.2.2 Vocal Usage Learning

Vocal usage learning, the capacity to produce specific calls in specific contexts as a result of experience, is a relatively common ability among mammals. Conditioning the production of a signal in response to a cue is a simple task to train and clearly demonstrates usage learning. Many animal species have been experimentally trained to vocalise when cued (for reviews see Adret, 1992; Janik & Slater, 1997; Schusterman, 2008; Seyfarth & Cheney, 2010) including: rats (*Rattus norvegicus*) (e.g. Lal, 1967), guinea pigs (*Cava porcellus*) (Burnstein & Wolff, 1967), dogs (*Canis lupus familiaris* (e.g. Salzinger & Waller, 1962), cats (*Felis catus*) (e.g. Molliver, 1963; Farley et al., 1992), dolphins (e.g. Lilly, 1965; Richards et al. 1984; Harley et al., 2005) and nonhuman primates (e.g. Myers et al., 1965; Randolph & Brooks, 1967; Wilson, 1975; Aitken & Wilson, 1979; Koda et al., 2007; Hage et al., 2013).

In these studies, the animals are typically trained similarly to a 'go/no go' paradigm; when given a particular cue the animal is trained to produce a specific sound, and when the cue is absent the animal remains silent, such that the behaviour is under stimulus control (Adret, 1992). Calls produced under

stimulus control can be made in novel contexts independent of behavioural and motivational states, clearly demonstrating the animal's ability to associate calls with new contexts. Training typically involves positive reinforcement such that the animal is rewarded, usually with food, for correct responses; however some studies have also utilized negative reinforcement such that aversive stimuli, usually electric shocks, are removed for correct responses (e.g. Lal, 1967).

Pinnipeds in particular have flexibility and control over call production. All three families have been trained to produce specific call types under stimulus control (e.g. Schusterman & Feinstein, 1965; Schusterman & Balliet, 1970; Schusterman et al., 1972; Schusterman, 1978; Shapiro et al., 2004; Schusterman & Reichmuth, 2008). In a notable example, visual and auditory sensitivity thresholds of the California sea lion were measured by training the animal to produce 'click' sounds upon seeing a target or hearing a tone (Schusterman & Balliet, 1970; Schusterman et al., 1972). Harbour seals and Pacific walruses are additionally capable of trained control over vocal plasticity; by varying positions of the mouth and tongue both species can produce novel sounds when cued (Schusterman & Reichmuth, 2008; Schusterman, 2008). This cannot be considered production learning as the animals did not conclusively produce these new calls as a result of experience with other sounds. Instead, such studies that have trained the animal to innovate (i.e. produce a novel sound each time the behavior is cued) demonstrate the animal can control when specific sounds are produced. Overall, pinnipeds appear to have control over when they produce sounds as well as the manipulation of their sounds' structure.

While it is anticipated that wild populations are also capable of vocal usage learning, it is difficult to document experimentally. Most evidence demonstrates the animals produce specific call types in specific context, but does not conclusively show this has arisen through learning; alternatively, call usage in these cases could be innately determined. However, usage learning is likely to play a role in these observed cases and appears to be primarily used for social communication. Calls can inform others of relevant external information, such as with food calls which are used by several species to

provide information such as the quality, quantity, and location of specific foods (e.g. Roush & Snowdon, 1994; Hauser, 1998; Clay et al., 2012; King & Janik, 2015). For example, adult cotton-top tamarins (*Saguinus oedipus*) produce two specific call types when given access to food (Roush & Snowdon, 1994). They produce these sounds almost exclusively in this context; these two call types are absent when food is not available, and usually are the only call types produced when food is available. Juvenile and sub-adult tamarins produce similar call types, however they produce them in a variety of contexts, not solely when food is available (Roush & Snowdon, 1994). This suggests that with age, animals may learn to use call types appropriately through experience, although other explanations (such as delayed maturation, or development of the nervous system) can be ruled out. Interestingly, in some of these cases, calls are used differently depending on audience; some call types are only used when addressing specific individuals that are of higher or lower social ranking, infants, or familial relations (e.g. Sherman, 1977; Cheney & Seyfarth, 1985; Owren et al., 1997).

Other potential examples of vocal usage learning in wild populations are predator alarm calls, specific calls which alert others to specific threats. For example, Campbells' monkeys (*Cercopithecus cambelli*) produce predator-specific alarm calls alerting conspecifics to the presence of hawk eagles (*Stephanoaetus coronatus*) or leopards (*Panthera pardus*) (Zuberbühler, 2001). For different predators different avoidance strategies are used (Zuberbühler, 2001). Alarm calls are used by several species (e.g. vervet monkeys, *Chlorocebus aethiops*: Seyfarth et al., 1980b; meerkats, *Suricata suricatta*: Manser et al., 2002; Hollén et al., 2006; red squirrels, *Tamiasciurus hudsonicus*: Greene & Meagher, 1998). In some cases these alarm calls are highly specific, indicating predator proximity, type and/or location (e.g. Seyfarth et al., 1980a & b; Manser et al., 2002).

Whether the use of predator alarm calls is learned and/or innate is of some debate. In a review, Hollén and Radford (2009) discussed that some research illustrates the role of learning in alarm call development, shown through the inappropriate use of and responses to alarm calls by young animals

(e.g. Mateo 1999; Hollén & Manser, 2007), while other research supports calls being innate, with young animals using and responding to alarm calls comparably to adults (e.g. Seyfarth & Cheney, 1980; Ivins & Smith, 1983). The ability to produce and withhold specific calls in such contexts impacts the animal's survival and evolutionary fitness. Given the high cost of inappropriate use and interpretation of these call types, it might be expected that these calls would be fixed from birth. However, Mateo (1996b) discussed that in some species a certain degree of plasticity in juvenile alarm calls would be valuable; if predators change over time, some flexibility could allow for better responses to be learned through experience. Overall alarm calls appear to be primarily innate and are later improved upon with experience.

2.2.3 Acoustic Comprehension Learning

Previously I discussed vocal usage learning, where animals learn to produce specific calls in specific contexts as a result of experience. This involves the animal acting as a signaller, and their calls provide information to others. Conversely, vocal comprehension learning allows for individuals to act as the receiver and interpret the meaning of the calls produced by others.

Comprehension learning is not limited to gaining information from conspecific calls. Here, I briefly discuss some examples of how animals may learn to associate biological sounds, from both conspecifics and heterospecifics, as well as non-biological sounds, such as anthropogenic signals, with external events to gain information. As this learning in not limited to vocalisations, I refer to this as acoustic comprehension learning, as it encompasses learning contextual significance of all sound signals.

Biological Sound Signals

Typically, comprehension learning refers to the ability of an individual to gain information from conspecific calls. As previously discussed with usage learning, this is relatively simple to train and demonstrate experimentally (e.g. Le Prell et al., 2002; Shapiro et al., 2004; Vergara, 2011). While most evidence cannot be conclusively attributed to learning, it is anticipated that wild

populations would be similarly capable of gaining information from conspecific signals, such as food calls (e.g. Hauser, 1998; Clay et al., 2012; King & Janik, 2015). They could also provide additional information about the behavioural state of the caller (e.g. Maros et al., 2008; Almonte, 2014; Keesom et al., 2015), or provide identity information. A unique example of this is the signature whistle of bottlenose dolphins; individual dolphins have distinctive frequency modulation patterns to one of their whistle types (e.g. Janik et al., 2006), which can be presumably associated with specific individuals.

Comprehension learning is not limited to conspecific calls. Animals can similarly gain information from human speech. Several species are able to understand hundreds of human words, including domestic dogs (Kaminski et al., 2004), nonhuman primates (Savage-Rumbaugh, 1986), sea lions (Schusterman et al., 2002), dolphins (Herman et al., 1993) and parrots (Pepperberg, 2002). In these cases, the animals were trained to respond differentially to specific spoken words through positive reinforcement. Such learning can thus be attributed to simple operant conditioning, associating a cued response with reward, which is not considered as evidence by most of the comprehension literature. However, as discussed by Seyfarth and Cheney (2010), the underlying mechanism by which animals learn associations between sound and an event are irrelevant as they do not change the signal's potential use for information. The way an individual learns how to respond to a conspecific call in comparison to a spoken word may vary, but both allow for receivers to gain information, clearly demonstrating comprehension learning. Although artificial in that the information is trained using conditioning techniques, trained responses provide equally compelling evidence for comprehension learning.

Individuals could also use comprehension learning to gain information by eaves-dropping on the calls of neighbouring individuals. Although most evidence cannot be conclusively attributed to learning, it is anticipated that wild populations utilize the ability. For example, eavesdropping on other species' communication signals is a widely used source of information, such as from the predator alarm calls of nearby species (e.g. Hauser, 1988; Seyfarth & Cheney,

1990; Fichtel, 2004; Rainey et al., 2004a & b) or the sounds produced by predators and/or prey (e.g. Hauser & Wrangham, 1990; Deecke et al., 2002). While it is possible that the cases may not be solely attributed to learning as the calls could have innate significance, experimental evidence supports the possibility. For example, the fringe-lipped bat, *Trachops cirrhosis*, eavesdrops on prey (frog) calls to identify edible from toxic species (Page & Ryan, 2006). This association between calls and food is learned; when a bat is exposed to reversed calls (i.e. the edible calls paired with the toxic species and vice versa) they learn to only approach the edible species. The learned association is also socially transmissible; if an inexperienced bat is paired with an experienced individual, they learn the reversed calls much faster (Page & Ryan, 2006). Predators can learn to use acoustic prey cues, and this information can be socially transmittable.

Anthropogenic Sound Signals

Arguably, any stimulus that is reliably associated with an external event could be used by animals for information. Thus, it would be expected that animals could also use non-biological sounds for information. This has been clearly demonstrated in trained contexts, such as for behavioural audiograms where animals have been rewarded for responding differentially upon hearing artificial tones (e.g. Schusterman et al., 1972; Wolski et al., 2003; Gaspard et al., 2012; Stansbury et al., 2014). Comparatively fewer studies have considered how wild animals may similarly use information provided by anthropogenic sounds. Instead, past studies have predominantly focused on the negative impacts of introduced noise (e.g. Nowacek et al., 2007; Tyack, 2009; Barber et al., 2010; Slabbekoorn et al., 2010; Tyack & Janik, 2013). Here, I focus the discussion on one example of how anthropogenic sounds could potentially be used, specifically by marine mammals, to gain information for prey detection.

Through comprehension learning, animals can associate specific sound stimuli with food availability. This would be most obvious where anthropogenic noise indicates prey patches. In the marine environment, anthropogenic signals such as from fishing boat engines, pingers, sonar and acoustic deterrent

devices could all be used by predators to locate prey. Marine mammals have been found to be attracted by such sounds (Chilvers & Corkeron, 2001; Thode et al., 2007), occasionally even to sounds introduced with the intention of deterring them (Bordino et al., 2002). In wild populations, higher incidences of predation at fisheries with acoustic deterrent devices (ADDs) may be attributed to learned associations between sound and prey (Jefferson & Curry, 1996). ADDs produce loud sounds that are believed to cause avoidance responses in species, such as seals, that predate upon fish farms. While seals that have not previously been exposed to ADD signals have been found to generally avoid them, seals that have experience finding fish at that location quickly habituate to ADD sounds (Götz & Janik, 2010 & 2013). Through operant conditioning, ADDs can be associated with the presence of fish, thus acting as a 'dinner bell' and potentially attracting predators to the area.

Current research utilizing artificial sound sources to mark fish (Cooke et al., 2011), could also be influenced by this 'dinner bell' effect. Many of these studies use acoustic coded transmitters (also known as pingers) to study fish populations and movements. These pingers typically emit an ultrasonic acoustic signal that is inaudible to fish, but is audible to some marine mammal predators (Bowles et al., 2010). If the signal is detectable, the sound could be associated with the presence of prey and cause increased predation. This appears to be a realistic concern; current studies have observed decreased survivorship rates for acoustically tagged juvenile salmon when compared to those with similar tags that produce no sound signal (Wargo-Rub et al., 2012a & b)

Most of the studies currently illustrating the use of anthropogenic sound as a signal for prey detection are opportunistic. It is currently unclear to what extent anthropogenic acoustic cues affect prey detection, or how long it would take for a predator to make an association between novel sound and an associated food source in natural environments. It is also important to consider behaviours are likely learned (such as the 'dinner bell effect') separately from behaviours that may be more physiologically or innately driven (such as the repellent qualities of ADD's). More controlled studies investigating the uses of anthropogenic sound for information are needed.

2.2.4   Vocal Production Learning

Introduction

As previously discussed, acoustic contextual learning (the usage and comprehension of sounds) is a relatively common ability across animal species. In comparison, vocal production learning has been well demonstrated in birds and humans, but among nonhuman mammals it is accepted as being relatively scarce. This possibly reflects a difference in the vocal capabilities among mammalian species; for example, thorough investigations of nonhuman primates have overall failed to demonstrate production learning (see Snowdon, 1990 for review).

Alternatively, there may be a lack of relevant studies for some species. Production learning is difficult to demonstrate when concrete evidence relies upon showing that animals have altered their repertoires outside of natural ranges. Few studies have been able to show such convincing evidence, and consequently production learning has only been conclusively documented in cetaceans (for reviews see Janik & Slater, 1997; Tyack & Sayigh, 1997), elephants (Poole et al., 2005; Stoeger et al., 2012), harbour seals (Ralls et al., 1985), and bats (Esser, 1994).

Here I will review current studies documenting vocal production learning. The majority of evidence comes from instances of call matching, or convergence, in which individuals produce new calls that are similar to a sound they have heard. First, I discuss the cases in which vocal production learning can be demonstrated in the strictest sense; where animals have shown clear changes in their repertoires which become more similar to a model that differs from species specific repertoires and cannot be explained alternatively. Such cases have provided conclusive evidence for the species currently accepted as production learners.

Additional evidence comes from cases of call imitation, in which animals imitate conspecific signals, but few of these studies can be considered as cases of true production learning as the calls may have pre-existed within an individual's repertoire, and thus observations could be attributed to learning to produce the calls in a new context (i.e. usage learning). However, with detailed

observations thoroughly documenting call repertoires both before and after exposure to models, particularly examining juvenile repertoires during over-production phases, imitation can provide convincing evidence that should be more widely considered as production learning. Conversely, animals may also produce calls that differ, or diverge, from acoustic models. Again, with detailed observations both before and after sound exposure, divergence is equally compelling evidence for vocal production learning.

I will then briefly review the literature supporting vocal learning in pinnipeds, such as the harbour seal that was anecdotally reported to imitate human speech (Ralls et al., 1985). Vocal learning is a complex task, and other capabilities are thought to be linked to production learning. The ability to physically move in time with a beat, referred to as motor entrainment, is thought to be tied to the neural adaptation for production learning (Patel, 2006). I will briefly discuss motor entrainment, and its relation to vocal learning. Then, the potential functions of vocal learning in mammals will be reviewed examining imitation, mimicry and how the ability may have evolved.

In summary, I will discuss how vocal production learning is currently viewed as a strict black or white capability, and suggest how future studies should reconsider this view when examining accepted vocal learners and evaluating different species' capabilities.

Call Convergence

The strongest evidence for production learning has come from the spontaneous and/or trained vocal mimicry of sounds not previously occurring in an individual's vocal repertoire. Mimicry of environmental sounds has been observed in an African elephant mimicking trucks (Poole et al., 2005), a harbour seal and elephants (Indian, *Elephas maximus indicus*, Asian, *Elephas maximus,* and African) mimicking human speech (Ralls et al., 1985; Stoeger et al., 2012), pilot whales (*Globicephala melas*) and false killer whales (*Pseudorca crassidens*) mimicking naval sonar (Alves et al., 2014; DeRuiter et al., 2013), a killer whale mimicking California sea lion barks (Foote et al. 2006), young bats mimicking computer generated calls (Esser, 1994), and bottlenose dolphins

mimicking trainer (Miksis et al., 2002) or computer generated whistles (Richards et al., 1984; Reiss & McCowan, 1993).

Vocal convergence, in which call structures become more similar within social groups (Tyack, 2008), is equally compelling evidence such as the whistle copying (Tyack, 1986; Smolker & Pepper, 1999; Janik, 2000; Watwood et al., 2004; Fripp et al., 2005) and matching (King et al., 2013) of bottlenose dolphins, killer whales (*Orcinus orca*) imitating dialect changes (Crance et al., 2014; Deecke et al., 2000), song convergence in humpback whales (*Megaptera novaeangliae*) (e.g. Payne et al., 1983; Guinee et al., 1983; Noad et al., 2000), an African elephant mimicking unique calls of Asian elephants (Poole et al., 2005), pygmy marmosets (*Cebuella pygmaea*) calls becoming more similar within mating pairs (Snowdon & Elowson, 1999), pups imitating their mothers echolocation in greater horseshoe bats (Jones & Ransome, 1993), and group calls becoming more similar with changing group composition in greater spear-nosed bats (*Phyllostomus hastatus*) (Boughman, 1998).

I believe these convergence cases conclusively demonstrate production learning as the observed call changes cannot be alternatively explained. Other studies have documented call convergence that could be attributed to production learning, but other explanations cannot be ruled out. For example, several studies suggest that multiple primate species may have group specific calls (e.g. Mitani & Gros-Louis, 1998; Marshall et al., 1999; Weiss et al., 2001; Crockford et al., 2004) and when introduced to new individuals and/or groups animals may converge to have calls similar to conspecifics (e.g. Snowdon, 1990; Elowson & Snowdon, 1994; Marshall et al., 1999; Seyfarth & Cheney, 2010; Watson et al., 2015). However there may be alternative explanations for these results. For example, rhesus monkeys produce a 'coo' vocalisation which was found to structurally vary in one observed matriline (Hauser, 1992). Hauser concluded that the differences observed were due to vocal matching among the familial group. However, the results could also be explained by genetics, in which group members were predisposed to produce similar calls due to morphological similarities.

Similarly, motivation may have accounted for some of these observed convergence cases. Several studies demonstrate that calls may be more variable when new groups or individual animals are introduced, and become more stereotyped over time (e.g. Snowdon & Elowson, 1999; Rukstalis et al., 2003). Animals exposed to unfamiliar individuals may produce more varied calls, which may become more stereotyped over time as those individuals become familiar, appearing to become more similar. Stress, such as might be expected from exposure to novel individuals, is known to alter many call parameters such as rate, duration, amplitude, frequency and modulation (Briefer & McElligott, 2012). With repeated exposures and experience, the stress of new individuals may be reduced and lead to call changes that could be mis-attributed to convergence. Additionally, some of these cases have shown that call variation within groups does not significantly differ. Convergence cannot be conclusively demonstrated if the calls fall within the group's normal variation, which has been argued when examining convergence of cross-fostered Japanese and rhesus monkeys (Masataka & Fujita, 1989; Musser et al., 2014).

Overall, cases of convergence provide convincing evidence for production learning, and as suggested by past reviews (e.g. Tyack, 2008) such subtle changes which occur within the species natural repertoire should be considered when examining vocal learning abilities. However, for production learning to be demonstrated in these cases, calls must be well documented both before and after changes occur, ideally measuring variability across the animal's entire repertoire. Documenting juvenile calls, which are likely to be overproduced, is ideal when comparing an individual's repertoire before and after learning from a model. Marler & Nelson (1992) argued that songbirds learn through 'action-based learning' such that during the plastic phase of development individuals overproduce song, and during crystallization specific songs are selectively socially reinforced, leading to certain songs being reproduced while others are discarded. By documenting repertoire composition during over-production phases, it is more likely that the full extent of an individual's repertoire is

measured. Additionally, call changes should be shown to vary between multiple groups, rather than looking at changes which occur in only one population.

Most of these studies have examined learning by documenting call frequency changes. However, other research has documented matching of other parameters. Bottlenose dolphins approximately match the number and duration of human speech sounds (Lilly, 1965). A beluga whale, "Noc" was reported to produce calls similar in amplitude modulation, intercall interval and harmonic structure to human speech (Ridgway et al., 2012). Although not produced by the larynx and thus not a traditional vocalisation, an orangutan (*Pongo pygmaeus x P. abelii*), "Bonnie" spontaneously produced 'whistles' similarly to a human model, and would match number and duration of examples (Wich et al., 2009). Animal communication systems are complex and variable across species, and although some species may not be capable of control over call frequency, they may have more control over parameters controlled by the respiratory system. These cases provide evidence for production learning in a wider range of species than have been previously considered. Further investigation into the production learning abilities of these acoustic parameters across species is warranted.

Call Divergence

The majority of evidence for production learning derives from studies of call matching. Production learning can also be convincingly demonstrated by vocal divergence, where call structures are changed to become more dissimilar to other sounds. For example, European free-tailed bats (*Tadarida teniotis*), adjust the frequencies of echolocation clicks to prevent overlap, or sound signal "jamming", with nearby conspecifics (Ulanovsky et al., 2004). When flying in groups of two or more the frequencies of each bat's calls shifted such that each individual had a unique call frequency. These calls were dynamic and frequency changed continually, such that individuals constantly adjusted to match changes and maintain individuality (Ulanovsky et al., 2004). Call divergence is also thought to allow for animals to compensate in noisy environments; for example beluga whales (*Delphinapterus leucas)* change call frequencies when exposed

to vessel noise (Lesage et al., 1999) and bats change the frequency of their echolocation clicks to make up for Doppler shifts created by differential velocity of the bat and its target (Schnitzler, 1973; Trappe & Schnitzler, 1982). Presumably, this allows for sound signals that might otherwise be masked to be perceived by the intended receiver. As discussed by Tyack (2008), although not usually considered in discussion of vocal learning, such call changes diverging from environmental noise are important to consider when examining a species' vocal learning capabilities. These cases of divergence conclusively demonstrate production learning, and should be more widely considered when evaluating vocal learning abilities,

Divergence may also provide additional support for vocal production learning in primates; in four communities of wild chimpanzees (*Pan troglodytes verus*) 'pant hoots' varied more between nearby communities than between those located farther away (Crockford et al., 2004). The authors found no significant differences in habitat or genetic relatedness of the groups, and concluded that calls were modified through learning (Crockford et al., 2004). Such geographic variation, where individuals have calls more similar within their groups than between groups, may be explained with production learning. However, these cases may also be alternatively explained by other factors, such as with social reinforcement (Janik & Slater, 1997).Overall, past reviews on production learning indicate that the ability may be common across a wider taxonomic range than currently accepted (Janik & Slater, 1997, 2000; Tyack, 2008). Careful, objective measurements documenting call changes converging and/or diverging from models (examining calls both before model exposure and changes occurring after) provide convincing evidence for production learning in a wider range of species (Tyack, 2008). There are several cases   which suggest several primate species are capable of subtle vocal production learning (e.g. Masataka & Fujita, 1989; Hauser, 1992; Mitani et al., 1992; Elowson & Snowden, 1994; Mitani & Gros-Louis, 1998; Crockford et al., 2004; Candiotti et al., 2012; Watson et al., 2015). Given such evidence further, more controlled research, on vocal learning in primates is warranted. Similar studies indicate the possible ability in Gunnison's prairie dogs (*Cynomys gunnisoni zuniensis*)

(Slobodchikoff & Coast, 1980), mice (Arriaga et al., 2012) and Rocky mountain pikas (*Ochotona princeps*) (Somers, 1973; Conner, 1982).

Vocal Learning in Pinnipeds

Of the mammalian species studied, production learning has been especially well documented in marine mammals, particularly in cetaceans. Investigations have only briefly examined pinniped species, but these few studies indicate the potential for production learning (see review by Reichmuth & Casey, 2014). California sea lions (Schusterman & Feinstein, 1965; Schusterman, 1978), Pacific walrus (Schusterman & Reichmuth, 2008), harbour seals (Ralls et al., 1985) and grey seals (Shapiro et al., 2004) are all capable of vocalizing under stimulus control, demonstrating usage learning. Additional control over vocalisations has been demonstrated in a harbour seal and four Pacific walruses, where the natural vocalisations were shaped to vary in amplitude, duration, frequency, modulation, and rate (Schusterman & Reichmuth, 2008; Schusterman, 2008). This culminated in training the animals to innovate and produce novel sounds on cue. The animals' production of novel sounds cannot be considered production learning because they may not have done so in response to exposure to other sounds. While it is possible the animals imitated novel sound models to produce new calls, it is more likely that shaping contingencies reinforced the production of varied call aspects in specific situations, illustrating usage learning and/or improvisation. However, the animal's ability to manipulate calls indicates sufficient physical control over their vocal tract to have the variability needed for production learning.

The strongest evidence for vocal learning in pinnipeds comes from a captive harbour seal, "Hoover", who was raised among humans and began to spontaneously emit vocalisations similar to speech (Ralls et al., 1985). These vocalisations were subsequently shaped using positive reinforcement to create a repertoire of several human words and phrases including "hello" and "come over here". A concurrent effort was made to train another harbour seal at the same captive facility to produce speech like sounds, but was only successful in shaping one call, the seal's name (Ralls et al., 1985). Another study similarly

attempted to teach a neonatal, male harbour seal, "Chimo", to copy speech and physical movements by raising the pup with human models; however this was also unsuccessful (Moore, 1996).

These studies examined vocal learning in captive pinnipeds, which may not reflect the species' natural abilities. Some evidence suggests that vocal learning may also occur in wild populations of elephant seals (Sanvito et al., 2007). Agonistic call types of juvenile and adult males in a breeding population were sampled over a period of eight years. The vocalisations of juveniles (three to five years old) were variable, and as they aged the calls became more stereotyped (Sanvito & Galimberti, 2000). Older males predominantly produced one call type, with each call type consisting of a sequence of pulsed calls that varied between individuals in temporal parameters including intercall and pulse duration (Sanvito et al., 2007). With age, the juveniles produced the call types of the dominant males (the most successful breeders) most often. The authors concluded that the young males imitated the vocalisations of the dominant adults (Sanvito et al., 2007). As the authors measured specific call types, it cannot be concluded that the seals were using production learning to copy calls. Rather, the call types could have previously existed in the juveniles' repertoire, and through usage learning they learned to preferentially produce the call types of dominant males during agonistic contexts.

This macrostructure approach suggests vocal learning had occurred by quantifying the proportion of call occurrence, but the results can be alternatively explained. The authors discounted the possibility that genetics favoured certain vocal types by comparing calls of six seals to those of their father. In these cases, none of the offspring produced the same vocal type as their fathers (Sanvito et al., 2007). The social status of the parent was not discussed; if the fathers were not dominant, this would support the study's results, but if the father was a dominant male the offspring should have imitated his vocalisations. Alternatively, conditioning could also explain the observed changes as certain call types could be socially reinforced differentially within the group.

Motor Entrainment

Other capabilities are thought to be tied to production learning. The vocal learning and synchronization hypothesis postulates that the ability to motor entrain to auditory stimuli (i.e. to physically move in time with an acoustic beat) is tied to the neural adaption for mimicry (Patel, 2006). Thus, the ability to move in time with a beat is thought to indicate the capability for production learning. This has been supported by the ability of established vocal learners to entrain, such as in cockatoos (*Cacatua galerita*) (Patel et al., 2009), African grey parrots (*Psittacus erithacus*) (Schachner et al., 2009) and budgerigars (*Melopsittacus undulates*) (Hasegawa et al., 2011), as well as the inability to find examples of entrainment in accepted non-vocal learners (Schachner et al., 2009).

However, some weak evidence suggests that two assumed non-vocal learners, rhesus monkeys and chimpanzees, can tap to a beat (Zarco et al., 2009; Hattori et al., 2013). Furthermore, conclusive evidence for motor entrainment has been recently demonstrated in California sea lions (Cook et al., 2013), which are assumed to not be vocal learners. These observations may indicate that motor entrainment has no relation with vocal learning. Alternatively, perhaps while all animals capable of vocal production learning are capable of motor entrainment, not all animals capable of motor entrainment are capable of vocal production learning. It is also possible that the assumption that these species are non-vocal learners may be false, and evidence of motor entrainment in a species may warrant further investigation for production learning abilities. It has been similarly argued that the ability for animals to imitate physical movements may also be linked to vocal imitation, and perhaps species capable of movement imitation should also warrant further testing for production learning abilities (Moore, 1992).

Functions of Vocal learning

Vocal production learning is a complex skill and is presumably costly. Having cost suggests there would be substantial benefit to individuals utilizing the ability; however there is debate as to what these benefits are and thus what the function of vocal production learning may be. The majority of evidence

exploring the functional role of vocal learning has come from the bird song literature, and while it has been widely investigated (e.g. Catchpole & Slater, 1995; Beecher & Brenowtiz, 2005; Kelley et al., 2008; Wickler, 2013) there are no theories which account for the many observed occurrences of vocal learning. This is in part due to the many differences observed between species and within individuals in song learning; for example, birds vary in when and how long they are receptive to a tutor's song (referred to as the 'sensitive period'), which tutors they will learn from, how accurately they reproduce the tutor's song, and size of the learned repertoire (Beecher & Brenowitz, 2005).

To better address the different functions of vocal production learning, the discussion has been separated into two biologically relevant categories; imitation and mimicry. While both involve individuals changing their calls to become more similar to other sounds, imitation is specifically copying sounds of conspecifics, while mimicry is copying any other sounds. While there are several hypotheses, this discussion is limited to those with the most evidence. Most studies have been conducted with birds, which will be the focus of the discussion, but mammalian examples will also be highlighted. Lastly, as the few species known to be capable of vocal production learning are diverse, I will briefly discuss the main theories for how the ability may have evolved in such distinct groups.

Imitation

Imitation, in which individuals change their calls to become more similar to conspecific models, is primarily thought to influence sexual selection in birds (Catchpole & Slater, 1995). Darwin (1871) originally proposed that bird song was used for inter-sexual selection, where male song influenced female mate choice. This hypothesis is supported as in existent taxa bird song is predominantly sexually dimorphic, with only males singing in most cases, suggesting female choice could act as a driving force for song evolution (Catchpole & Slater, 1995). There is evidence for females choosing both mates with songs more similar to those she experienced in early life, presumably as a way of ensuring compatibility (Nottebohm, 1972; Riebel, 2003), as well as those

with different songs, presumably to avoid inbreeding (Grant & Grant, 1996). Song learning is also thought to serve as an honest indicator of male quality as repertoire size and accuracy is thought to be developmentally costly (Nowicki et al., 1998; Nowicki & Searcy, 2005). In some species, females have been found to prefer males with larger repertoires (e.g. Nowicki et al., 2000; Searcy & Yasukawa, 1996) and songs that more closely match tutors (Nowicki et al., 2002).

However, a recent investigation indicates that such inter-sexual selection pressures may not solely explain the evolution of song (Odom et al., 2014). In a retrospective survey, Odom et al. (2014) found that in the ancestors of many songbirds females sang. As singing was not ancestrally exclusive to males, this leads to questions concerning the extent to which inter-sexual selection explains song evolution. While inter-sexual selection is likely to still play a role, other forms of selection (as discussed below) are important to consider.

Song imitation also plays a role in mediating aggressive interactions through intra-sexual selection in which typically males compete for access to females (e.g. Brown & Farabaugh, 1997; O'Loghlen & Rothstein, 2012). Song can also be used to maintain territories. Territorial songbirds have been observed to 'counter-sing' when song from one bird is responded to with a song from another bird. This appears to serve a role in identifying known neighbours from strangers, referred to as the 'dear-enemy' effect (Fisher, 1954). Using a neighbour-stranger playback technique first developed by Weedon and Falls (1959), several bird species display fewer aggressive behaviours (such as wing flapping or increased song rate) when played a known neighbour's song in comparison to a stranger's song (for a review see Falls, 1992). However, this is only true when the neighbours' song is played from that neighbours' normal territory; if the neighbour's song comes from a novel location the bird responds similarly to that of a stranger (Falls & Brooks, 1975). This appears to mediate aggressive territorial interactions; for previously known neighbours there would be no advantage in wasting energy on already established territories. For a stranger the cost of strongly responding would presumably be an investment in maintaining existing territory (Catchpole & Slater, 1995).

Further playbacks have investigated the responses of song sparrows (*Melospiza melodia*) to song type match (in which the same song is imitated), a repertoire match (another song apart of both birds repertoire is produced), or an unshared song (Burt et al., 2001). While stranger (i.e. unshared) song resulted in the most aggressive behaviours, type matching resulted in more aggressive behaviours than repertoire matching (Burt et al., 2001). Type and repertoire matching thus appear to mediate neighbour conflicts as disputes can be escalated or de-escalated depending on the type of match used (Burt et al., 2001).

Such song exchanges could also allow birds to assess the distance of rivals (Morton, 1996). Interestingly, chipping sparrows (*Spizella passerina*) appear to use song when forming temporary alliances against intruders; when a stranger enters a bird's territory, neighbouring males will join him in simultaneous defensive singing displays (Goodwin & Podos, 2014). Such uses of song are presumably advantageous as they would allow for the avoidance of conflict and potentially increase breeding success by reducing time and energy cost (Payne, 1982, 1983).

While inter- /intra-sexual selection appear to be the main functions of bird song imitation, there are many additional theories for its evolution. Song could be adapted to habitat, where song aspects that transmit most efficiently within a habitat are learned, such that over generations song will become best suited to a bird's habitat (Hansen, 1979). Alternatively, song learning may allow for specific song features to maximize transmission while minimizing costs, such as preventing damage to the vocal system (Nottebohm, 1991).

While the functions of vocal imitation have been mostly examined in birds, strong evidence is available for one mammal; the dolphin. As previously discussed, dolphins produce individually distinctive calls known as signature whistles (e.g. Caldwell, 1965; Caldwell et al., 1990). These can be used to identify individuals using differences in frequency modulation, even when individual voice characteristics are removed (Janik et al., 2006). When isolated from group members signature whistles are the most commonly emitted call type of an individual's whistle repertoire when isolated from group members

(e.g. Caldwell et al., 1990). As dolphins predominantly produce signature whistles when group members are separated, and much less frequently when the group is together, they are thought to function as a cohesion call to maintain group contact (Janik & Slater, 1998). Although infrequent, dolphins have also been observed to copy and match signature whistles of other group members in both captive and wild populations (Janik & Slater, 1998; Janik, 2000; Quick & Janik, 2012; King et al., 2013). As signature whistle matches occur predominantly between close social relations, such as mother-calf pairs and male alliance partners, they are thought to be an affiliative signal for maintaining social bonds (King et al., 2013). Dolphins will also respond to playbacks of their own signature whistle by calling back with the same whistle, further supporting that the whistle serves as an identity label when communicating within a group (King & Janik, 2013). While signature whistles are relatively stable over an individual's life (Sayigh et al., 1990), whistles have been observed to change with social relationships. Juvenile male alliance partners have been shown to change their signature whistles to become more similar to their partner (Watwood et al., 2004). In these cases, signature whistles could also provide information about social relationships in addition to individual identity.

Additional evidence for the function of vocal learning has been found in bats. In some species, pups will match the frequency parameters of their mother's calls. This could serve a social function, such as for mother-pup bonding and recognition, or be used to assist mate-selection by preventing inbreeding (e.g. Esser & Schmidt, 1989; Knörnschild, 2014). In other cases, rather than changing calls to converge with those of other individuals, bats diverge to make their calls distinctive to avoid masking. As previously discussed, individual European free-tailed bats change the frequencies of echolocation calls when in groups, presumably to avoid 'jamming' of signals used by multiple animals. (Ulanovsky et al., 2004).

Mimicry

The function of vocal mimicry, in which individuals change their calls to become more similar to a model other than conspecifics, is unknown (see

reviews by Baylis, 1982; Kelley et al., 2008; Wickler, 2013). This may in part be due to the wide use of the term 'mimicry' within the literature. Historically, Bates (1862) introduced the term 'mimicry' to describe the evolutionary deception he observed in butterflies and insects. Different species, some of which were edible and some toxic to predators, had similar bright coloration patterns. Bates presumed that predators avoided both the edible and inedible species as they were deceived and could not distinguish the honest signals from the mimics (Bates, 1862). This mimicry can be attributed to evolutionary convergence, as species with similar natural selection pressures evolved to have similar traits (Kelley et al., 2008).

However, here the discussion is limited to 'learned mimicry', in which individuals change to become more similar to a model within their lifetime (rather than across generations). For learned mimicry to occur, this model cannot be a conspecific (otherwise learning would be deemed imitation). However, this presents another grey area within the terminology; take for example a situation in which an individual learned a call from a different species model. Thereafter, a conspecific learns the same call, but from a member of its own species. While the learned sound is the same, the first example would technically be mimicry, and the second imitation (Baylis, 1982; Kelley et al., 2008).

Baylis (1982) and Kelley et al. (2008) have provided excellent reviews of the different theories for the function of learned vocal mimicry within birds, which are briefly discussed here. As done by Kelley et al. (2008), the theories are separated by the intended receiver; interspecific communication, in which an acoustic cue is received by other species, and intraspecific communication, in which the acoustic cue is received by conspecifics (Kelley et al., 2008).

Interspecific vocal mimicry (mimicry of non-conspecific sounds where the intended receiver is heterospecific) is presumed to deter predators or competitors. Batesian acoustic mimicry could be used by a caller to deceive a potential threat by mimicking species that predate upon that threat (Dobkin, 1979). For example, subjective observations of great bowerbirds (*Chlamydera nuchalis*) have reported the birds making sounds similar to other local predators

(cats, dogs and kites) to presumably intimidate approaching humans by imitating other predatory threats (Frith & Frith, 2004). Similarly, a caller could mimic to attract another species, presumably to mob or attack the immediate threat. For example, if the caller mimicked wounded prey, other predators would pursue the mimicked prey and give the caller the opportunity to escape (Curio, 1978; Hogstedt, 1983). Mimicry could also facilitate brood parasitism, in which one species lays eggs in another species nest to be raised by host parents. Vocal mimicry could reduce the chance of rejecting the parasitic nestling by the host parents (Davies, 2011).

Intraspecific vocal mimicry (mimicry of non-conspecific sounds where the intended receiver is conspecific) is presumed to facilitate mate selection. Sexual selection could favour males which mimic if there is a cost associated with learning larger repertoires, and thus males with the largest repertoires are preferred mates by females (Loffredo & Borgia, 1986; Nowicki et al., 2002). Similarly, females could prefer novelty in song repertoires, and males which mimic other species may be selected for variety rather than quality of matches (ten Cate & Bateson, 1988).

Despite these many possible theories to explain vocal mimicry, Kelley et al. (2008) concluded that there was very little evidence to support any hypothesis. Instead, they posit that most instances of vocal mimicry can be explained by a 'learning mistakes hypothesis' in which mimicry is the result of errors. Birds may learn other sounds as a by-product of the ability for imitation, and thus mimicry serves no real purpose (Kelley et al., 2008). This is supported as simple, common sounds which are closer to species-specific repertoires are more typically mimicked, suggesting that birds learning conspecific calls may have inadvertently learned other sounds (Kelley et al., 2008).

Evolution

The capacity to imitate or mimic sound is a comparatively rare ability in the animal kingdom, with accepted vocal learners being limited to humans, birds, cetaceans, bats, elephants and phocid seals (e.g. Janik & Slater, 1997; Tyack, 2008). Within each of these groups, accepted vocal learning species are closely

related to non-vocal learning species, raising questions about the evolutionary phylogeny for vocal learning. Given that the ability has developed in such different groups, but is not common to all of the species within the groups, it is hypothesized that the capability evolved independently from common ancestors (Nottebohm, 1972; Jarvis, 2004; Jarvis, 2006). While in some species vocal learning may have been selected for, in others it may have been selected against, resulting in some closely related species being capable while others are not (Jarvis, 2006).

Jarvis (2006) reviewed the hypotheses which may explain why vocal learning would be selected for; vocal learning could be advantageous for individual identification, semantic communication, territorial defence, mate attraction, or allow callers to adapt sound properties for better propagation in changing environments. He also discussed that once the ability evolved within a taxonomic group, it may then be selected against in particular species. Given the hypothesis that a varied call repertoire is attractive to potential mates, it might similarly be expected to be attractive to predators. Thus, particularly in prey species, vocal learning would be more likely selected against. This is partially supported given that of the accepted vocal learning mammalian species, many appear to be top predators (Jarvis, 2006).

However, this hypothesis assumes that prey species with small, fixed repertoires would be less vulnerable to predators than those with large repertoires. Jarvis (2006) attributes this to habituation; species with fixed repertoires would frequently produce the same calls, and due to over exposure predators would learn not to attend to the signal. This seems unlikely; in learning paradigms, variably rewarded responses result in increased occurrence of behaviour (Ferster & Skinner, 1957). In this case, a predator occasionally using a sound signal to obtain prey would be expected to increasingly use the signal for information, depending on how often the signal leads to a successful prey capture. It might also be expected that if fixed sound signals are used by specific prey, predators may be more likely to use the signal as it is a reliable indicator of prey location. If their prey produced variable sound

signals, predators may not be able to use them as reliable indicators, especially if the varied sounds could be confused with other, non-prey species.

Do Production Learning Abilities Fall on a Continuum?

The process by which individuals modify sounds from learned experience is complicated, and currently there are many questions concerning mammalian vocal learning. Given the ability has arisen in very different animal groups, but is absent in closely related species, its evolution is of particular interest. While some hypothesize that selection for and against may explain its origins, it is also possible that the ability is more common than has been previously accepted.

In humans and many other animal species, sound production can be explained by source-filter theory (e.g. Fant, 1960; Titze, 1994). In its simplest form, sound is produced when air is released from the lungs into the larynx, causing the vocal folds to vibrate and generate a sound wave. Thus, the larynx is the sound's source. The sound consists of a fundamental frequency, determined by the vibration rate, and higher frequency components, called harmonics. These harmonics are integer multiples of the fundamental frequency. As the sound travels from the larynx through the vocal tract some of these frequencies are emphasized or suppressed. Emphasized resonant frequencies are referred to as formants, and although they often correspond to harmonics, they are not limited to harmonic frequencies. The vocal tract thus serves as the sound's filter. By changing the shape of cavities within the vocal tract (for example by moving the tongue or changing the shape of the lips), the amplitude of different frequencies are changed (e.g. Fant, 1960; Titze, 1994).

The currently accepted vocal learning species have all displayed arguably advanced control over vocal production by varying the frequency of their calls by manipulation of the sound's source and/or filter. Although some species may not be capable of control over call frequency, they may have control over a less complicated aspect of the vocal tract, the respiratory system, which controls sound parameters such as amplitude, duration and call rate. While currently accepted non-vocal learning species may not be able to control the source or filter, they may be capable of manipulating sound parameters controlled by the

respiratory system which have not been considered in most of the vocal production literature (Janik & Slater, 1997).

The current black or white classification of species as either vocal learners or non-vocal learners may underestimate production learning abilities. Vocal learning may be better measured across species using a continuum, considering the ability to control the sound's source and filter separately from the respiratory system. Further investigation into the vocal production learning abilities of these parameters is warranted.

Similarly, all of the currently accepted vocal learning species have mimicked calls outside of species specific repertoires. To truly show production learning, a call cannot have pre-existed in the animals' repertoire. Many of the conclusive demonstrations of production learning have been striking examples of animals mimicking abnormal sounds such as human speech, trucks, and computer signals (Ralls et al., 1985; Esser, 1994; Poole et al., 2005; Stoeger et al., 2012). However, should we expect that all production learners will match such atypical signals? When considering evidence of imitation, convergence and divergence, in conjunction with examining parameters controlled by the respiratory system, a larger range of species appear to be capable of production learning (notably including some primate species). As suggested by past reviews (Janik & Slater, 1997, 2000; Tyack, 2008), re-evaluation of the currently accepted vocal learning species and further investigation into vocal production learning considering subtler call changes is warranted.

2.3   Vocal Development in Mammals
2.3.1   Introduction

Age-related changes in an individual's call structure can be attributed to physical development, maturation of the nervous system and/or learning (Sanvito et al., 2008). However, it is difficult to separate how these processes individually affect call ontogeny as all presumably play critical roles in an individual's repertoire development. In this section, I will review studies documenting vocal development and discuss the changes that are most likely attributed to physical maturation.

It is expected that morphologically determined acoustic parameters would change in relation to physical development (though not necessarily with a linear trend). These sound features should differ between groups of individuals, such that animals from similar age, weight and sex (particularly in sexually dimorphic species) classes should have similar call parameters. Within these groups, inter- and intra-individual variation should occur within fixed ranges as morphological constraints would restrict variability. These changes should also occur comparably both in normal individuals and those exposed to 'abnormal' acoustic environments, such as cross-fostered or deafened individuals.

As humans develop, the vocal tract increases in length and the position of larynx within the vocal tract descends (Fitch & Giedd, 1999), resulting in developmental changes to sound parameters. This is thought to occur similarly in most other mammals (e.g. Fitch, 2006; Taylor & Reby, 2010). During development, young animals increase in size and weight. Increased body weight is believed to correlate with increased length of the vocal folds, changing the vibration rate and decreasing fundamental frequency (Titze, 1994). Body weight also correlates with vocal tract length, with increased length correlating with decreased call frequencies (Fitch, 1997). Thus, as mammals age, there should be a negative correlation between size and call frequency parameters. Similarly, size should positively correlate with maximum call duration and amplitude (Fitch & Hauser, 2003). Larger individuals have increased lung capacity, which would allow for production of longer, louder sounds.

Developmental changes are also likely to differ between sexes, particularly in sexually dimorphic species. Some of these differences can be attributed to size, as in many species males and females grow to different sizes at different rates. Call changes could also be hormonally mediated. In humans, sex hormones affect laryngeal development and elasticity of the vocal folds (Abitbol et al., 1999). Such changes would alter the vibration rate, and thus change the frequency of calls. Additionally, some species develop sex specific structures for sound production, such that some signals can only be produced by one sex (Balaban, 1994).

2.3.2   Overview

Throughout the rest of this section, the current literature on vocal development will be examined considering changes that can most likely be attributed to physical maturation. Call ontogeny has been widely documented in birds, particularly in songbirds due to their striking vocal learning capabilities (Catchpole & Slater, 1995), which will be briefly discussed. Comparatively fewer investigations have examined how calls develop in mammals. I will then review studies concerning vocal development in mammals by mammalian order, and, where appropriate, class.

The studies included in this review have all been peer-reviewed and examine objectively measured changes in temporal, frequency and/or energy parameters (see Table 2.2 for a summary of cited studies). Many studies were excluded as they focused exclusively on call type changes over time. As different studies may classify call types differently, type-based investigations are difficult to compare and they have been excluded from this review. Additionally, several studies have been conducted and published in other languages. Due to the logistics of translating, works that have not been published in English could not be evaluated, and were excluded from this review.

Parameter measurements varied greatly between studies, which makes comparison across species difficult. Throughout this review, care was taken to discuss parameter changes as they were reported by individual study to avoid confusion. However, if multiple parameters changed in a similar trend then generic labels were used to describe changes. For example, if multiple frequency parameters (such as minimum, maximum, fundamental and peak frequency) all changed similarly these were discussed in general terms as frequency changes.

From this review, I identify developmental trends across species. Most mammals are anticipated to follow source-filter theory predictions, where frequency will negatively correlate and amplitude and duration positively correlate with age and size. Similar trends would be expected within sexes but should differ between sexes, particularly in sexually dimorphic species.

Table 2.1. Summary of the studies cited in this review. Symbols used: **Age** h=hour, d=day, wk=week, m=month, yr=year, Neo=neonate, Inf=Infant, Juv=Juvenile, Ad=Adult, **Sex** ♂=Male, ♀=Female, **Social Experience** W=wild, C=captive, I=isolated, MO=mother only, FO=family only, SG=social group **Call Parameters** ↑= increased with age, ↓= decreased with age, Freq=frequency, Fund=fundamental, Min=minimum, Max=maximum, Mo=modulation, Dur=duration, #=number, RR= repetition rate, Inf=inflections, Amp=amplitude, BW=bandwidth, Harm=harmonic, ICI=interclick interval, Var=variability. NA= not available. * = studies which examined developmental changes of specific call types.

Common Name	Genus, Species	Age range	Sex	Social Exp.	Summary of Call Changes	Source
CETACEA						
Bottlenose Dolphin	*Tursiops truncatus*	Birth – 1 yr	7 ♂, 1 ♀	C, SG	Whistle Freq Mo ↑ up to 5 – 8 m, stabilized by 1 yr.	McCowan & Reiss, 1995*
Bottlenose Dolphin	*Tursiops truncatus*	Birth – 7 wk & 1 yr	1 ♂	C, MO	Echolocation click train Dur and # clicks ↑, while max and mean click RR ↓. Click trains stabilized 7 wk – 1 yr.	Favaro et al., 2013*
Bottlenose Dolphin	*Tursiops truncatus*	Birth – 5 d	1 ♂	C, SG	Calf calls were longer in Dur and lower in Freq than adult calls.	Killebrew et al., 2001*
Bottlenose Dolphin	*Tursiops truncatus*	Birth – Ad (>13 yr)	9 ♂, 12 ♀	W & C, SG, MO & I	Sound loops, Freq Mo, and Dur ↑. Variability ↓.	Caldwell & Caldwell, 1979*
Bottlenose Dolphin	*Tursiops truncatus*	Birth – 50 hr & 771 hr	1 ♂, 1 ♀	C, SG	Dur and # Inf ↑. Freq ↑ until 8 hr old, then ↓.	Morisaka et al., 2005*
Bottlenose Dolphin	*Tursiops truncatus*	Birth – 3 & 11 m	2 ♀	C, SG	Infant clicks were lower in Freq in comparison to adult 'clicks'.	Manoukian et al., 2002*
Bottlenose Dolphin	*Tursiops truncatus*	2, 7, & 38 wk	1 ♀	C, SG	Echolocation click peak Freq ↑.	Lindhard, 1988*
Bottlenose Dolphin	*Tursiops truncatus*	Birth – 40 d	2 ♂	C, SG	Freq, Freq Mo, and Dur ↑. At 3-4 wks calves showed open mouth echolocation, by 5 wks closed mouth echolocation.	Reiss, 1988*
River Dolphin	*Neophocaena phocaenoides*	Birth – 181 d	1 ♂	C, SG	Produced clicks by 22 d, comparable in Freq and Dur to adults. Click RR, train Dur and # 'clicks' ↓.	Li et al., 2007*
Beluga Whale	*Delphinapterus leucas*	Birth – 3 yr	1 ♂	C, MO & SG	Pulsed call Amp, RR, Freq, and Freq BW ↑. Whistle Amp, Freq, and FM ↑. Dur, variability and # Inf ↓.	Vergara & Barrett - Lennard, 2008*
Sperm Whale	*Physeter macrocephalus*	Neo (< 1 m)	1 ♂, 1 ♀	C, I	Calf clicks had low directionality and centroid frequency, and longer duration in comparison to adults.	Madsen et al., 2003*
Sperm Whale	*Physeter macrocephalus*	4-8 meters (size proxy of age)	2 ♂, 2 ♀	C & W, SG	Calf clicks were more variable and had lower amplitude in comparison to adults.	Watkins et al., 1988*

Common Name	Genus, Species	Age range	Sex	Social Exp.	Summary of Call Changes	Source
CARNIVORA:	**PINNIPEDIA**					
Southern Elephant Seal	*Mirounga leonine*	Juv (< 8 yr) & Ad (> 8 yr)	129 ♂	W, SG	Juveniles produced more calls with ↑ dur and var. Formants (especially higher Freq) ↓ from 7-8 yrs old	Sanvito et al., 2008*
Mediterranean Monk Seal	*Monachus monachus*	1 wk – 5 m	1♂, 2♀* (captive)	W & C, MO & SG	Freq ↑ for all pups, Dur of ♀ calls ↑.	Muñoz et al., 2011*
Hawaiian Monk Seal	*Monachus schauinslandi*	Birth – 21 d	NA	W, SG	Dur ↑, Fund Freq and Peak Harm Freq ↓.	Job et al., 1995
Leopard Seal	*Hydrurga leptonyx*	Sub-Ad (1-3 yr), Ad (>4yr)	10 ♂	W, SG	Young animals produced more calls per bout, ↑ bout Dur, and more variable Freq than adults.	Rogers, 2007*
Bearded Seal	*Erignathus barbatus*	3 wk – 6 yr	3 ♂, 3 ♀	C, SG	Only ♂ vocalised, and only started producing calls upon reaching sexual maturity. Dur and Freq ↑.	Davies et al., 2006*
Harp Seal	*Phoca groenlandica*	Birth – 15 d	49 ♂, 42 ♀	W, SG	There were no significant differences in calls by age, however peak harm freq discriminated between sexes.	Van Opzeeland & van Parijs, 2004
Harbour Seal	*Phoca vitulina*	1 – 42 d	NA	C, SG	Freq and Freq Mo ↓. Dur, # of Infl, and mean Freq discriminated between sexes.	Khan et al., 2006
Subantarctic Fur Seal	*Arctocephalus tropicalis*	Birth – 7 m	NA	W, SG	Fund Freq and emphasized Harm Freq ↑. Freq Mo ↓ during the first 15 days.	Charrier et al., 2003
CARNIVORA:	**CANIDAE**					
Swift Fox	*Vulpes velox*	Approx. 3 wk – 5 m	NA	C, SG & FO	Fund Freq ↓, BW and Dur ↑ for noisy calls. However changes were not consistent across call types.	Darden & Dabelsteen, 2006*
African Wild Dog	*Lycaon pictus*	3 – 7 wk	NA	W, FO & SG	Fund Freq ↓ for some call types.	Robbins & McCreery, 2003*
Timber Wolf	*Canis lupus*	Birth – 6 wk	4 ♂, 1 ♀	C, FO	Fund Freq ↓, majority of change occurs over first 5 days, after which continues to ↓ at slower rate.	Coscia et al., 1991*
Domestic Dog	*Canis familiaris*	Birth – 4 wk / 6 m	NA	C, SG	Chihuahua calls ↑ in variability 1st 6 d, by 10 d ↑ Freq and Dur. Calls similar to adult by 4 wks. Irish Setter/Doberman ↑ Freq	Cohen & Fox, 1976*

Common Name	Genus, Species	Age range	Sex	Social Exp.	Summary of Call Changes	Source
CARNIVORA:	**FELIDAE**					
Domestic Cat	*Felis catus*	Birth – 170 d	5 ♂, 5 ♀	C, SG & MO	Fund and Peak Freq ↑ to day 11, then ↓. Dur ↑. Freq ranges, upper Freq limit ↓ 19 d, then stable.	Romand & Ehret, 1984*
Domestic Cat	*Felis catus*	30 d – 3 yr	5 ♂, 5 ♀	C, SG, FO & MO	Freq ↓ throughout the 3 yrs.	Shipley et al., 1988
5 Big Cat Species	Genera *Panthera* & *Puma*	Varied, All < 1 yr	NA	C, NA	Peak Freq ↓. Freq differs between *Panthera* & *Puma,* attributed to the larynx position.	Peters, 2011*
CARNIVORA:	**URSIDAE**					
Giant Panda	*Ailuropoda melanoleuca*	6 – 21 yr	9 ♂, 9 ♀	C, Unk	Jitter ↑, Min Fund and difference in formants ↓. ♂ ↑ Freq Mo and jitter, ↓ difference in formants than ♀	Charlton et al., 2009*
CHIROPTERA						
Lesser Spear-nosed Bat	*Phyllostomus discolour*	Birth – 50 d	NA	C, I	Freq Mo ↑, while the Min Freq of the lowest modulation point ↓.	Esser, 1994*
Seba's Short-tailed Bat	*Carollia perspicillata*	Birth – 7 wk	NA	C, MO	Peak Freq ↑ and Dur ↓. By day 50 calls were comparable to adult echolocation in Dur and Freq.	Sterbing, 2002*
Jamaican Fruit Bat	*Artibeus jamaicensis*	1 d – Ad (> 104 d)	NA	C, SG	Freq BW, # calls, and sweep rate ↑, Dur and ICI ↓.	Carter et al., 2014*
Little Brown Bat	*Myotis lucifugus*	Birth – 6 wk	NA	C, MO	Freq and Freq Mo ↑, Dur ↓.	Moss et al., 1997*
Big Brown Bat	*Eptesicus fuscus*	1 d – 5 wk	3 ♂, 6 ♀	C, FO	Freq ↑, Dur and variability ↓. By 4 wks calls were comparable to adult echolocation.	Monroy et al., 2011*
Big Brown Bat	*Eptesicus fuscus*	21-55 d, Ad* (age undefined)	NA	W & C, MO	Freq and pulse RR was lower in young compared to adults.	Gould et al., 1981*
Big Brown Bat	*Eptesicus fuscus*	Birth – 7wk	NA	NA	Freq ↑, Dur and variability ↓. Between 21-28 d calls were comparable to adult echolocation.	Moss, 1988*
Mustached Bat	*Pteronotus parnellii*	1 d – 5 wk	16 ♂, 12 ♀	C, SG	Freq and Amp ↑, Dur and variability ↓ for frequency modulated calls. For constant frequency calls Dur ↑.	Vater et al., 2003*
Japanese House Bat	*Pipistrellus abramus*	Birth – 1 m	3 ♂, 2 ♀	C, MO	Freq ↑ and Dur ↓. By 28 d calls were comparable to adult echolocation in Dur and Freq.	Hiryu & Riquimaroux, 2011*
Lesser Bulldog Bat	*Noctilio albiventris*	Birth – 4 m, Juv & Ad (age undefined)	NA	W & C, SG	For constant frequency calls Freq ↑. Produced more constant frequency and fewer frequency modulated signals with age.	Brown et al., 1983*

Common Name	Genus, Species	Age range	Sex	Social Exp.	Summary of Call Changes	Source
CHIROPTERA	Continued					
Schneider's Leaf-nosed Bat	*Hipposideros speoris*	Birth – 55 d	NA	W & C, SG	Fund and Peak Freq, and RR ↑, Dur and variability ↓. Young bats call through mouth, adults produce nasal calls.	Habersetzer & Marimuthu, 1986*
Flat Headed Bats	*Tylonycteris pachypus & T. robustula*	Birth – 25 d	10 ♂, 8 ♀	C, SG	Freq ↑ and Dur ↓.	Zhang et al., 2005*
Rufous Horseshoe Bat	*Rhinolophus rouxi*	1 – 42 days (size estimate)	NA	W, I* (orphaned)	Freq ↑. Young bats call through open mouth, adults produce nasal sounds. By 4-5 wk, produced adult echolocation calls.	Rübsamen, 1987*
Rufous Horseshoe Bat	*Rhinolophus rouxi*	3 – 4 wk	11 ♂, 4 ♀	W & C, SG & I	Freq ↑.	Rübsamen & Schäfer, 1990*
Greater Horseshoe Bat	*Rhinolophus ferrumequinum*	5 – 7 wk, 1 – 28 yr	NA	W, SG	Freq ↑ increased most up to 1-2 yrs., and then at a slower rate from 2-3 yr. Freq was stable, then ↓ from 10-23 yrs.	Jones & Ransome, 1993*
Greater Horseshoe Bat	*Rhinolophus ferrumequinum*	1 – 4 wk	NA	W & C, SG, MO & I	Fund and Peak Freq ↑, variability ↓. Young bats call through mouth, adult's nasal sounds. By 4 wks adult echolocation	Matsumura, 1979*
Black Flying Fox	*Pteropus alecto*	1 – 35 d (size estimate)	9 ♂, 12 ♀	W & C, SG, MO & I	There were no significant differences in calls by age or sex, however there was individual variability in Dur and Peak Freq.	Van Parijs & Corkeron, 2002*
PRIMATE						
De Brazza's Monkey	*Cercopithecus neglectus*	Juv.(1–3/4) & Ad. (> 3/4)	8 ♂, 15 ♀	C, FO	Freq and Dur variability ↓, ♀ had higher Freq call than ♂. Juveniles and ♀ had higher call rates than adult ♂.	Bouchet et al., 2012*
Baboon	*Papio ursinus*	Juv, Sub-Ad & Ad (age undefined)	12 ♂, 9 ♀	W, SG	Mean Freq and formant dispersion ↓, Dur for some call types ↑. Smaller, younger animals had shorter Dur and ↑ Fund Freq.	Fischer et al., 2002*
Cape Baboon	*Papio ursinus*	Juv (< 2 yr), Sub-Ad (4-6) & Ad (>6 yr)	21 ♂, 37 ♀	W, SG	Dur ↑, Fund and Peak Freq ↓ (in ♂ more so than ♀). Freq Mo ↓.	Ey et al., 2007*
Aye Aye	*Daubentonia madagascariensis*	Inf & Ad (age undefined)	2 ♂, 2 ♀	W & C, SG, MO & I	Freq and Dur ↓.	Stanger & Macedonia, 1994

Common Name	Genus, Species	Age range	Sex	Social Exp.	Summary of Call Changes	Source
PRIMATE	Continued					
Rhesus Macaque	*Macaca mulatta*	1 wk – 5 m	NA	C, SG & MO	Freq and Freq Mo ↓, Dur ↑.	Hammerschmidt et al., 2000*
Japanese Monkey	*Macaca fuscata*	Inf (< 1 yr), Juv (1-3 yr), & Ad (>4 yr)	14 ♂, 36 ♀	W, SG	Fund Freq and variation in Max Fund Freq ↓.	Inoue, 1988
Squirrel Monkey	*Saimiri sciureus*	Birth – 2 yr	NA	C, SG & MO	Freq ↑, noise and Dur ↓.	Hammerschmidt et al., 2001*
Squirrel Monkey	*Saimiri sciureus*	Birth – 6/17 m	2 ♂, 5 ♀	C, MO & I	There were no significant differences by age. Infant calls were comparable to adult repertoires.	Winter et al., 1973*
Vervet Monkey	*Cercopithecu aethiops*	Inf(<10m),Juv(10-18m), Ad (>4yr)	23 Ad ♀ Inf NA	W, SG	Fund and Peak Freq, and variability (both Freq and Amp) ↓, and intercall interval of pulsed calls ↑.	Hauser, 1989*
Pygmy Marmoset	*Cebuella pygmae*	Birth – 2 yr	6 ♂, 2 ♀	C, FO	Dur and Freq Mo ↑, and Freq, variability in Freq Mo, and Min Freq ↓ in most litters.	Elowson et al., 1992*
Common Marmoset	*Callithrix jacchus*	3 – 25 wk	NA	C, FO	Call rate ↓. Changes in calls varied by call type.	Pistorio et al., 2006*
Ringtailed Lemur	*Lemur catta*	Inf (<1 yr), Juv (1-2), Ad (>2 yr)	NA	C, SG	Freq appears to ↑.	Macedonia, 1993*
PROBOSCIDEA						
African Elephant	*Loxodonta africana*	Neo (< 1 m) – 18 m	6 ♂, 5 ♀	C, SG	Fund Freq ↓ and Dur ↑, especially with low-frequency 'rumbles'.	Stoeger-Horwath et al., 2007*
ARTIODACTYLA						
Goitred Gazelle	*Gazella subgutturosa*	1 – 24 wk	10 ♂, 13 ♀	C, SG	Freq ↓. Dur ↑.	Efremova et al., 2011*
Goitred Gazelle	*Gazella subgutturosa*	3 – 6 & 23 – 26 wk	18 ♂, 17 ♀	C, SG	Fund, formant and Peak Freq ↓.	Lapshina et al., 2012*
Red Deer	*Cervus elephas*	2 – 4 d, 2 – 4 yr	8 Ad ♀, 7 Unk	C, MO	Freq ↓	Vaňková & Mάlek, 1997
Red Deer	*Cervus elephas*	Sub Ad (5 – 6 yr) & Ad (7– 3 yr)	57 ♂	W, SG	Max Fund & Formant Freq ↓. Formant Freq similar correlated to stag weight	Reby & McComb, 2003*
Roe Deer	*Capreolus capreolus*	Varied by sex	39 ♂, 9 ♀	W, SG	Freq ↓. ♀ ↑ Amp of High Freq than ♂, and age variations did not mask sex.	Reby et al., 1999*
Goat	*Capra hircus*	Birth – 124 d	18 ♂, 5 ♀	C, MO & FO	Freq, Freq & Amp Mo ↓. In ♀, Dur, Freq and Freq Mo ↑, jitter/shimmer ↓, Fund more stable than ♂.	Briefer & McElligott, 2011*
Beef Cattle	*Bos taurus*	Opportunistic, birth-180 days	14 unk.	C, SG	Formant freq ↓, fund freq constant. Age related changes differed by individual, but not sex.	Padilla de la Torre et al., 2015*

Common Name	Genus, species	Age range	Sex	Social Exp.	Summary of Call Changes	Source
DASYUROMORPHIA						
Northern Quoll	*Dasyurus hallucatus*	18 - 109 d	NA	NA	Young began to vocalise between 35-45 days. Freq ↓, and became similar to adults by 109 days.	Aitkin et al., 1996
RODENTIA						
Albino mouse	*Mus musculus*	1 - 12 d	NA	C, FO	Dur, Freq BW and Amp ↓. Freq changes were variable between individuals.	Noirot & Pye, 1969
Woodmouse	*Apodemus sylvaticus L.*	Approx. 1-15d	NA	C, FO	Call rates and # of bouts ↑ over the first wk, then ↓. Approach adult rates around third wk.	Pontet et al., 1989
Ground Squirrel	*Spermophilus suslicus, S. fulvus*	Juv (< 1 yr) & Ad (> 1 yr)	NA	W, SQ	There were no significant differences between age groups.	Matrosova et al., 2007*
Guinea Pig	*Cavia porcellus*	2 – 40 d	20 ♀	C, FO	Dur ↓	Arch-Tirado et al., 2000*
5 vole Species	*Microtus*	Birth - 10 d	NA	C, FO	Intercall Dur and Fund Freq ↓, Fund Freq BW ↑.	Colvin, 1973*
6 rodent Species	*Various*	Approx. 3-21 d	22 ♂, 22 ♀	C, FO	Pups vocalised less often with age (↓ between 3-15 days).	Motomura et al., 2002*

2.3.3 Literature Survey: Aves

Vocal development in birds has been thoroughly investigated, presumably due to their relative ease to keep in captivity, quick maturation, and advanced vocal learning abilities. Several comprehensive reviews have examined call ontogeny across avian species (e.g. Nottebohm, 1970; Kroodsma, 1982; Catchpole & Slater, 1995; Baptista & Gaunt, 1997). While the majority of research has focused on the effects of learning on call development, here I will briefly discuss the changes most likely attributed to physiological development.

As it can be difficult to separate the effects of learning from physical maturation, most information on maturational call changes come from non-vocal learning bird species and/or abnormal individuals such as those cross-fostered or deafened. If these individuals develop normal adult repertoires, then call changes can presumably be attributed to morphological development (e.g. Lade & Thorpe, 1964; Nottebohm & Nottebohm, 1971). These studies indicate that in general calls decrease in frequency and increase in duration and amplitude with age, but differ between sexes (e.g. Cosens, 1981; Ballintijn & ten Cate, 1997; Klenova et al., 2014).

Repertoire changes in songbirds appear to be predominantly learned, and due to extreme flexibility of the sound production system are not thought to be restricted by many physiological limitations (Podos, 1996). As many birds learn the songs of conspecifics, presumably individuals would be physically capable of reproducing the tutor songs. However, physiological development can affect call memorization and trigger production. In many songbirds, singing is exclusive to sexually mature males, following an auditory template model for development (Catchpole & Slater, 1995). Young birds start with basic calls from birth. They go through a memorization phase, during which time they are exposed to model songs but do not begin to sing. Memorization is restricted to a 'sensitive period', a developmental time window during which individuals are capable of learning song models. Song production is initiated during the motor phase, when the males reach sexual maturity and testosterone triggers song production (Catchpole & Slater, 1995). Initial songs are variable, going through 'sub-song' periods where individuals use auditory feedback and duets with

models before developing crystalized, adult song. While this occurs in many species, other species are open-ended learners that can learn new calls throughout their lives, such as parrots (e.g. Pepperberg, 1994).

In existent species, bird song appears to be predominantly used by males and the ability is thought to play a significant role in female mate selection. While typically females do not sing, in some species birds with abnormally elevated testosterone levels produce song similarly to males (Garamszegi et al., 2007). Closer examination shows that in several species females are capable of producing learned song, and that in ancestral species females may have regularly sang (Odom et al., 2013). Song learning appears to occur in both sexes, but given the sex differences in production, in many cases it appears to be mediated by hormonal states.

2.3.4 Literature Survey: Mammalia

Cetacea

The vocal repertoire of many cetacean species has been relatively well investigated (for a review see Richardson et al., 1995). Comparatively few of these studies have investigated call development. The majority of information on cetacean vocal ontogeny concerns bottlenose dolphins, presumably due to their prevalence in captivity. Bottlenose dolphins produce broadband (clicks and burst pulses) and narrowband (whistles) sounds (e.g. Janik, 1999). Although relatively little is known about the use of burst pulses, clicks are primarily used for echolocation and whistles for social communication. The majority of whistles produced by a dolphin are signature whistles; a unique frequency modulation pattern which encodes individual identity information and is used to maintain group cohesion (e.g. Janik et al., 2006). Dolphins have also been observed to occasionally match the signature whistles of other group members, usually those they have a close social relationship with, suggesting signature whistle matching may be an affiliative signal facilitating social bonding (Janik & Slater, 1998; Janik, 2000; King et al. 2013).

Based on observations of captive bottlenose dolphins, calves produce burst pulses from birth (Morisaka et al., 2005). Calf echolocation clicks develop

by increasing in peak frequency, click train duration and number of clicks, and decreasing in repetition rate (Lindhard, 1988; Reiss, 1988; Manoukian et al., 2002; Favaro et al., 2013). Calves begin to click within the first two weeks (Lindhard, 1988; Reiss, 1988; Favaro et al., 2013). Calves begin to echolocate using an open mouth between three to four weeks, and progress to closed mouth echolocation at five weeks (Reiss, 1988). However, other observations have found calves do not begin echolocate until two to five months after birth (Manoukian et al., 2002). Calf clicks stabilize and resemble adult echolocation trains within their first year, perhaps as early as seven weeks (Favaro et al., 2013).

Whistles appear to develop by increasing in peak frequency, frequency modulation, number of loops, and duration (Caldwell & Caldwell, 1979; Reiss, 1988; McCowan & Reiss, 1995; Killebrew et al., 2001; Morisaka et al., 2005). As whistles become more complex in frequency modulation with age, they become less variable and develop into stereotyped whistle types (Caldwell & Caldwell, 1979; McCowan & Reiss, 1995). Whistles have been observed from birth (Morisaka et al., 2005) or within the first week (Killebrew et al., 2001), and begin to stabilize within the first five to eight months, resembling adult repertoires within the first year (McCowan & Reiss, 1995).

The vocal development of a few other cetacean species have been documented. Echolocation clicks of a neonatal Yangtze river dolphin (*Neophocaena phocaenoides*) were found to be comparable to adult clicks in frequency and duration (Li et al., 2007). A captive, male calf was monitored from birth and began to produce echolocation at 22 days old. These first clicks were comparable in frequency and duration to adult echolocation, but repetition rate, train duration, and number of clicks decreased with age until 112 days old, at which time they stabilized and were indistinguishable from adult click trains (Li et al., 2007).

Call development of a captive, male beluga whale (*Delphinapterus leucas*) calf, "Tuvaq", was monitored from birth to three years (Vergara & Barrett-Lennard, 2008). At birth, only low energy, broadband pulse trains were produced. These trains increased in energy, peak frequency, frequency

bandwidth, and repetition rate with age. The calf began to whistle at two weeks, which increased in energy, peak frequency, and frequency modulation with age. At four months the calf began to mix call types and produced pulse trains and whistles together. Tuvaq's calls resembled that of adult repertoires within the first year, however his calls continued to change after this time. At birth, he was housed only with his mother, but after one year was housed in a group including his father. At 18 months, Tuvaq incorporated a new call type into his repertoire that resembled one of his father's calls (Vergara & Barrett-Lennard, 2008). The addition of this call type to Tuvaq's repertoire was attributed to learning, however the authors acknowledge that it is also possible that genetics or delayed maturation could explain the calls development (Vergara & Barrett-Lennard, 2008).

Clicks from neonatal sperm whale (*Physeter macrocephalus*) calves have been opportunistically recorded from rehabilitation animals and during boat follows (Watkins et al., 1988; Madsen et al., 2003). Calf clicks were more variable, with lower amplitude, directionality, centroid frequency and longer durations in comparison to adults. Although these results should be interpreted with caution as rehabilitation animals may have abnormal repertoires, they appear consistent with the few wild recordings documented (Watkins et al., 1988; Madsen et al., 2003).

Although these studies provide a basic understanding of developmental trends in cetaceans, caution should be taken when interpreting these findings due to methodological limitations. For example, older studies were limited by technological developments and many did not record the full frequency range of cetacean calls (e.g. Reiss, 1988). Additionally, it can be difficult to identify which animal produces specific sounds; for example, McCowan and Reiss (1995) distinguished callers using bubble streams, which has been since found to be a unreliable indicator biased towards certain call types (Fripp, 2005). As each study recorded animals with variable regularity and rates, it is also difficult to compare when calls develop. One study found that bottlenose dolphin calves do not begin echolocate until two to five months after birth (Manoukian et al., 2002), while others within the first few weeks (Reiss, 1988; Lindhard, 1998;

Favaro et al., 2013). Conversely, an unpublished master's thesis documented echolocation clicks occurring much earlier, within 24 hours of birth (Malinka, 2014). Given variable recording schedules, it is difficult to make conclusions as to when certain call changes occur during development. Although some trends are apparent in cetacean call development, further investigation is warranted and interpretation should be made cautiously.

Carnivora: Pinnipedia

While the repertoires of many pinniped species have been well documented (for a review see Richardson et al., 1995), comparatively little is known about developmental changes. Most studies on call ontogeny have investigated phocid seals (Job et al., 1995; van Opzeeland & van Parijs, 2004; Davies et al., 2006; Khan et al., 2006; Rogers, 2007; Sanvito et al., 2008; Muñoz et al., 2011). Only one study has examined an otariid species, the sub-antarctic fur seal (Charrier et al., 2003). Across species there are varying trends in frequency changes. With age, frequency increased in Mediterranean monk seals (*Monachus monachus)*, bearded seals (*Erignathus barbatus)*, and sub-antarctic fur seals (Muñoz et al., 2011; Davies et al., 2006; Charrier et al., 2003) but decreased in Southern elephant seals (*Mirounga leonina*), Hawaiian monk seals (*Monachus schauinslandi)*, and harbour seals (Sanvito et al., 2008; Job et al., 1995; Khan et al., 2006). However, in leopard seals (*Hydrurga leptonyx*) there were no changes in frequency with age (Rogers, 2007). Across species call duration, rate, and frequency modulation appear to decrease with age (Job et al., 1995; Charrier et al., 2003; Khan et al., 2006; Rogers, 2007; Sanvito et al., 2008; Muñoz et al., 2011). Calls may also decrease in variability, and become more stereotyped with age (Rogers, 2007; Sanvito et al., 2008).

Call development also differs by sex. In several species male and female calls appear to differ in peak frequency and duration (van Opzeeland & van Parijs, 2004; Khan et al., 2006; Muñoz et al., 2011). Calling rate and usage of some call types may also differ between sexes. Davies et al. (2006) observed the call development of captive bearded seals after weaning for six years.

Notably, the females never vocalised and the males only began to call upon reaching sexual maturity (Davies et al., 2006).

Carnivora: Canidae

A few studies have investigated age-related call changes in canines. Peak and fundamental frequency appear to increase over the first four weeks in domestic dogs (*Canis familiaris*) (Cohen & Fox, 1976). In swift foxes (*Vulpes velox*) and African wild dogs (*Lycaon pictus*), pups older than three weeks decreased call frequency parameters (Robbins & McCreery, 2003; Darden & Dabelsteen, 2006). In the foxes, the frequency bandwidth and duration of broadband, noisy calls increased with age (Darden & Dabelsteen, 2006). However, these parameter changes were not consistent across call types, and pups predominantly produced different call types at different stages of development. This may be due to motivational states; younger pups may produce calls to elicit maternal attention, while older pups which are no longer as dependent on care may shift to produce calls for social interactions, such as conspecific aggression (Tembrock, 1976). Such a change in call type and repertoire use may explain why domestic dog pup calls increased in frequency during the first few weeks of life, when they were most dependent on maternal care, and swift fox and African wild dog pups decreased in frequency after three weeks, when the pups' dependency on maternal care decreased.

Carnivora: Felidae

Vocal ontogeny of felines has been thoroughly examined in domestic cats (*Felis catus*). Normal hearing kittens raised with their mothers increased call fundamental and peak frequency for the first 11 days, at which point they begin to decrease (Romand & Ehret, 1984). Calls continued to decrease in frequency and increase in duration for three years, correlating with physical growth through maturation (Shipley et al., 1988). However the upper frequency limit, and thus frequency range, of kitten calls decreased only over the first three weeks (Romand & Ehret, 1984).

Normal call development has been compared to abnormal individuals in deafened kittens and those without adult models. Deafened kittens followed similar developmental trends, but produced more variable, longer duration, higher amplitude calls with lower frequencies compared to normal hearing animals (Romand & Ehret, 1984; Shipley et al., 1988). Although the deafened kittens had overall lower frequency calls, they followed similar frequency changes with age. Isolated kittens, which were not exposed to any adult cats including their mothers, produced calls that were longer in duration than normal hearing animals (Romand & Ehret, 1984). As overall frequency changes occurred in normal and abnormal kittens, they are most likely attributed to morphological development. As deafened kittens produced louder and longer calls with more variability, amplitude and call consistency is presumably controlled by auditory feedback. Isolated kittens housed away from mothers produced longer duration calls, which suggest that call duration is influenced by behavioural motivation, presumably as these kittens were attempting to elicit maternal care (Romand & Ehret, 1984).

Similar results were found in big cat species, where frequency changes were negatively correlated with body weight (Peters, 2011). Frequency and weight changes in the first year of life were examined in five big cat species; lions (*Panthera leo*), jaguars (*Panthera onca*), leopards (*Panthera* pardus), tigers (*Panthera tigris*) and pumas (*Puma concolor)*. All five species had relatively stable weight gain throughout the year, and their mean dominant frequency of calls decreased consistently as weight increased. In *Panthera* species, the larynx descends and the vocal tract lengthens with growth. The dominant frequency changes in *Panthera* calls occurred in formant frequencies. In the *Puma*, the larynx does not descend and the vocal tract remains relative short in comparison to *Panthera*. The frequency changes in *Puma* occurred in the fundamental frequency, and were considerably higher frequency than in *Panthera*. These results support the source-filter theory of mammalian sound production, and indicate that in *Panthera* dominant frequency changes are produced within the vocal tract (i.e. the filter) while in *Puma* they are produced by the larynx (i.e. the source) (Peters, 2011).

Carnivora: Ursidae

Age-related repertoire changes of one bear species has been examined in adult animals (Charlton et al., 2009). Male and female giant panda (*Ailuropoda melanoleuca*) calls were compared between animals ranging from 6 to 21 years old. Older animal calls had more jitter (mean difference in changes of the fundamental frequency across the call) and frequency modulation. These changes were attributed to a potential decrease in vocal fold mass and elasticity with age (Titze, 1994). There were also sex differences between calls, with males having increased jitter and frequency modulation and less difference between formant frequencies than females, potentially due to sexual dimorphism as males are larger than females. The frequency difference between formants was also found to decrease with age, but only in females. This was presumed to be mediated by hormonal differences between the sexes; higher levels of oestrogen are correlated with increased stiffness of the vocal folds. Changes in the female panda's oestrogen levels with age were thought to be related with the observed formant frequency changes (Charlton et al., 2009).

Chiroptera

Vocal development in microchiropteran bats is of interest given their echolocation and vocal learning abilities (Jones & Ransome, 1993; Ulanovsky et al., 2004; Knörnschild, 2014). Several studies have documented the development of echolocation (for a review see Gould, 1975). From birth, bats produce constant frequency (CF) and frequency modulated (FM) sounds (Brown et al., 1983). New born bats primarily produce low frequency FM 'isolation' calls, which presumably serve a mother-attraction function (e.g. Habersetzer & Marimuthu, 1986). As the pups age, this 'isolation' call appears to increase in frequency and decrease in duration, evolving into typical adult echolocation clicks within the first two months of life (e.g. Moss, 1988; Sterbing, 2002; Monroy et al., 2011; Carter et al., 2014). While clicks normally do not appear within the first week of life, pups can be induced to produce clicks from birth with movement (Vater et al., 2003). Normally, pups remain stationary for

the first few weeks of life. However, if placed on a pillow and moved vertically, pups as young as an hour old were observed to produce low frequency clicks (Vater et al., 2003).

The frequency increase in pup calls appears to correspond with the development of the auditory system (Rübsamen, 1987). In Rufous horseshoe bats (*Rhinolophus rouxi*), pups predominantly produced low frequency (10-15 kHz), pure tone 'isolation' calls in the first few weeks of life. The onset of hearing occurred by the second week, when the pup's hearing range was 15-45 kHz (peak sensitivity 15-25 kHz). In the third week, pups begin to produce echolocation clicks, with increased peak frequency (30-60 kHz) that corresponded with increased upper frequency hearing limit (57-60 kHz). By the fourth to fifth week, pups produced adult-like echolocation clicks (peak frequency 70-80 kHz) and obtained adult hearing (upper frequency limit 80 kHz) (Rübsamen, 1987).

While the majority of these studies examined call changes in the first few months of life, frequency parameters appear to change over the bats lifetime (Jones & Ransome, 1993). Greater horseshoe bats (*Rhinolophus ferrumequinum*) were followed for up to 28 years. Echolocation peak frequency increased the most within the first two years, and then at a slower rate through the third year. Call frequency was then stable until 10 years of age, at which time frequency slowly decreased until the bats were 23 years old (Jones & Ransome, 1993). This change could potentially correspond with age-related changes in hearing; high frequency sensitivity decreases with age in many mammals (e.g. Li & Borg, 1991; Ridgway & Carder, 1993). As the pups' call frequency increased corresponding with increased upper frequency hearing, potentially a presbycusal decrease in high frequency hearing could similarly correspond with a decrease in vocal frequency. However, there are currently no known audiograms documenting age related hearing changes in adult bats for comparison to call frequencies. It is also possible that the age related frequency decrease could be attributed to changes in the call production systems, such as a decrease in elasticity of the vocal folds. Such physical changes, both in

hearing sensitivity and the call production systems, could both contribute to changes in vocalisations.

Other call parameters appear to change with age. Amplitude of CF and FM calls and modulation of FM calls increased with age (Esser, 1994; Moss et al., 1997; Vater et al., 2003). Infants also produced more variable calls, appearing to go through 'babbling' stages comparable to humans (Knörnschild et al., 2006). Call variability decreased with age as more stereotyped sounds developed (Matsumura, 1979; Habersetzer & Marimuthu, 1986; Moss, 1988; Vater et al., 2003; Monroy et al. 2011).

Although bats are capable of vocal learning (see review by Knörnschild, 2014), many call changes appear to be attributed to morphological development as infants raised in abnormal conditions all follow the same developmental pattern and attain species-typical adult repertoires. Infant bats that have been deafened (Woolf, 1974), raised without adult models (Gould, 1975) and in helium-oxygen atmospheres (Gould et al., 1981) and all developed normal adult calls. While overall changes in echolocation appear to be attributed to morphological constraints, some call features appear to be determined by learning, such as previously discussed with infant bats matching the peak frequency of their mothers' calls (for review see Knörnschild, 2014).

While most investigations have focused on microchiropteran species, vocal development in one megachiropteran species, the Black flying fox (*Pteropus alecto*), has been investigated (van Parijs & Corkeron, 2002). 21 pups were recorded between birth and 35 days old (age was approximated using the infant's forearm length). There were no significant differences in calls by age or sex; however there was individual variability in duration and peak frequency.

The developmental differences between the calls of micro- and megachiropteran species may be attributed to call function; microchiropteran species use echolocation for navigation and prey detection, and would presumably need to develop functional echolocation before their first flight. Flying foxes do not echolocate, and would not have this pressure. Maternal care strategies also differ between the groups. In microchiropteran bats, pups are

typically cared for between four to six weeks, at which point they begin to fly, echolocate and forage independently. In comparison, black flying fox pups receive maternal care for approximately four months (Nelson, 1965). The pups are highly dependent upon their mothers for the first three weeks of life, during which time they continually nurse (Nelson, 1965). At approximately one month, the mother begins to leave to forage before returning to her young (Nelson, 1965). Potentially, pups maintain stable, individually unique calls throughout this period to aid mother-pup recognition (van Parijs & Corkeron, 2002). If pups maintain stable calls to facilitate individual recognition, ontogenetic call changes may be delayed during this extended nursing period.

Primates

Vocal development in primates is of particular interest due to their evolutionary relatedness to humans. Many studies have examined primate call ontogeny to better understand human vocal development, despite notable differences in vocal learning capabilities (e.g. Green, 1981; Snowdon & Hausberger, 1997; Corewyn, 2003; Ey et al., 2007; Hammerschmidt & Fischer, 2008).

Primates have extensive parental-care compared to many other mammals, with many species developing over several years before reaching adulthood (Altmann, 1987). In many species peak and/or fundamental frequency of calls decreases with age (Inoue, 1988; Hauser, 1989; Elowson et al., 1992; Stanger & Macedonia, 1994; Hammerschmidt et al., 2000; Fischer et al., 2002; Ey et al., 2007), although a few increase in frequency, such as the squirrel monkey (*Saimiri sciureus*) (Hammerschmidt et al., 2001) and ring-tailed lemur (*Lemur catta*) (Macedonia, 1993). Duration of calls predominantly increased with age (Elowson et al., 1992; Hammerschmidt, 2000; Fischer et al., 2002; Ey et al., 2007), but was found to decrease in Aye Aye's (*Daubentonia madagascariensis*) (Stanger & Macedonia, 1994). Overall, frequency modulation, variability, and noise decreased with age (Inoue, 1988; Hauser, 1989; Hammerschmidt et al., 2000; Hammerschmidt, 2001; Ey et al., 2007;

Bouchet et al., 2012). Additionally, intercall interval of pulsed calls was found to increase with age in vervet monkeys (*Cercopithecu aethiops*) (Hauser, 1989).

However, there are notable differences between studies. Winter et al. (1973) found no significant differences in squirrel monkey calls by age, and concluded that infant calls were comparable to adult repertoires. This differs from Hammerschmidt et al.'s (2001) findings that infant squirrel monkey calls were higher in frequency and lower in noise and duration compared to adult calls. This discrepancy may be due to differences in sound recording and analysis technology. Early efforts to document call changes were limited by the sensitivity of their equipment, while more recent endeavours can more acutely measure call changes.

Several studies have documented age related changes exclusively by examining changes in call types and their usage in primates. These were excluded from the discussion as comparison between studies is difficult when call types may be classified differently. However, examining call development by call type does appear to be relevant as in some studies parameter changes appear to vary between call types. In common marmosets (*Callithrix jacchus*) recorded between 3 and 25 weeks of age call changes varied for different call types (Pistorio et al., 2006). The infants' calls were classified into four call types; 1) the 'phee', a periodic, long duration call, 2) the 'trillphee', similar to a 'phee' with sinusoidal amplitude and frequency modulation at the start of the call, 3) the 'trill' which is sinusoidally amplitude and frequency modulated throughout the entire call, and 4) the 'twitter', which consists of multiple increasing frequency modulated sweeps. Across call types, calling rate decreased with age. In 'phee' and 'twitter' calls, frequency decreased. Frequency changes were variable across individuals for 'trillphee' calls. Duration of 'phee' calls and frequency modulation of 'trill' calls decreased (Pistorio et al., 2006). Overall, the decrease in frequency across call types suggests that changes in frequency may be morphologically determined, while changes that varied across types may be learned or determined by motivational states.

Call development in primates also varies with sex. In De Brazza's monkeys (*Cercopithecus neglectus*) and cape baboons (*Papio ursinus*), infant

calls were comparable between sexes. As they aged, female call frequency decreased at a slower rate, resulting in adult females having higher call frequencies than males (Ey et al., 2007; Bouchet et al., 2012). This is most likely attributed to animal size as females are smaller than males. Similarly, smaller and younger baboons produced shorter duration calls with higher fundamental frequencies compared to larger, older animals (Fischer et al., 2002).

Overall, primate calls appear to be relatively fixed from birth, and in species in which changes do occur they are most likely attributed to morphological maturation. Infant primates that have been cross-fostered (Owren et al., 1993; Seyfarth & Cheney, 1997; Hammerschmidt et al., 2000), deafened (Winter et al., 1973; Hammerschmidt et al., 2001) or raised without adult models (Winter et al., 1973; Hammerschmidt et al., 2001) develop species-typical adult repertoires. While there is some evidence that the fundamental frequencies of cross-fostered rhesus (*Macaca mulatta*) and Japanese macaques (*Macaca fuscata*) differed (Masataka & Fujita, 1989), there was no significant difference in fundamental frequency between the two species. The difference in cross-fostered infant development could thus be attributed to normal species-specific development (Owren et al., 1993).

Proboscidea

Vocal ontogeny of elephants is of particular interest given their complex social structure and vocal learning capabilities (Poole et al., 2005). Repertoire development of one species, the African elephant (*Loxodonta africana*), has been documented (Stoeger-Horwath et al., 2007). Two calves, a male and a female, were captive born at the Vienna zoo and recorded regularly with their family group from birth through eighteen months. Nine additional calves, five males and four females, were recorded opportunistically at an orphanage facility where they were kept with other calves and raised by humans. As the elephants aged, fundamental frequency decreased and duration increased, particularly in some call types such as low-frequency 'rumbles'. These changes were consistent for both the calves kept with their family in the zoo and orphaned

animals, suggesting the changes could be attributed to physical development as opposed to learning from adult models. Generally, there were no significant differences between sexes, with the exception that orphaned males vocalised more frequently than females in a feeding context (Stoeger-Horwath et al., 2007).

Artiodactyla

Vocal development in ungulates is of particular interest as some species have a relatively low positioned larynx within the vocal tract, which is more similar to human anatomy than in other mammals (e.g. Fitch & Reby, 2001). In some species such as the red (*Cervus elephas)* and fallow deer (*Dama dama*), males are capable of voluntarily dropping the larynx's position further towards the sternum. Dropping the larynx provides additional control over the sounds produced, and would allow for the production of lower call frequencies (Fitch & Reby, 2001).

Overall, ungulate calls decrease in fundamental, peak and formant frequencies and modulation, and increase duration with age (Vaňková & Málek, 1997; Reby et al., 1999; Reby & McComb, 2003; Efremova et al., 2011, Briefer & McElligott, 2011; Lapshina et al., 2012). In beef cattle (*Bos taurus*), formant frequencies and the difference between formants decreased with age, while the fundamental frequency remained constant (Padilla de la Torre et al., 2015). In several species, these changes in frequency and duration were also correlated with increased animal weight (Reby & McComb, 2003; Briefer & McElligott, 2011). Despite the potential for dishonest signalling given the animals' ability to drop the larynx and further decrease frequency parameters, overall the calls appear to serve as honest signals reliably indicating size (Reby & McComb, 2003). Call frequency and duration could also be used to indicate individual identity and sex in addition to size (Reby et al., 1999; Briefer & McElligott, 2011; Lapshina et al., 2012).

Dasyuromorphia

Among mammals, marsupials are interesting as the young spend little time developing in utero. The majority of neonatal development occurs in the mother's pouch. Vocal development has been observed in one marsupial, the Northern quoll (*Dasyurus hallucatus*) (Aitkin et al., 1996). 38 pups were recorded between 18 and 109 days old. Pup hearing was also measured using ABR and startle response methods. Young quolls began to vocalise between 35 and 45 days, when they started to leave their mothers teats for short periods of time. As the pups aged, frequency decreased, with calls becoming comparable to adult sounds by 109 days. Pup hearing increased in sensitivity and upper frequency limit with age, reaching adult sensitivity between 75 and 81 days (Aitkin et al., 1996).

Rodentia

Vocal development in rodents has been widely examined as a proxy to measure neurological and behavioural development (Branchi et al., 2001). In lab species including mice (*Mus musculus*) and rats, young pups produce periodic, ultrasonic calls that are influenced by environmental factors such as temperature (e.g. Allin & Banks, 1971; Sales & Smith, 1978; Mandelli & Sales, 2004) and chemical exposure (e.g. Hård et al., 1988; Carden & Hofer, 1990; Hofer, 1996). Normal pup sounds typically increase in call rate and number of bouts over the first week of development and then decrease in rate, becoming more similar to adult calls within the first month (e.g. Pontet et al., 1989; Motomura et al., 2002). Manipulation of the pups, both in utero and neonatally, has been found to significantly change call rate and provides a reliable indicator of abnormal neuro-behavioural development (Hofer, 1996).

In other rodent species pup call duration, amplitude and frequency decreases with age (e.g. Noirot & Pye, 1969; Colvin, 1973; Arch-Tirado et al., 2000), and fundamental frequency bandwidth increases with age (Colvin 1973). However, many of these changes were not documented in the majority of the literature as several of these studies occurred before advances in sound recording and analysis techniques became widely available, which limited the

ability to monitor such parameter changes. The majority of research investigating call changes in rodents, including abnormal call development, relied upon documenting changes in call rates and duration.

A more recent study investigating the alarm calls of ground squirrels (*Spermophilus suslicus & fulvus*) found no significant change in frequency parameters with age or size, contradictory to source-filter theory predictions (Matrosova et al., 2007). The results were attributed potentially to imitation, where juvenile squirrels imitated the calls of adults to reduce predation. Juvenile ground squirrels are vulnerable to predators such as polecats and weasels that do not prey on adults. Thus, there may be an evolutionary pressure for juveniles to have calls similar to adults. However, the authors were unable to explain the physical mechanism that would allow young animals to imitate calls of older animals (Matrosova et al., 2007).

2.3.5   Trends in Vocal Development

In this section I will examine the developmental trends that occur across species. In general, most mammalian calls decrease in frequency and increase in duration and amplitude parameters with age. Of the 73 species reviewed, 40 followed this trend, which is consistent with source-filter theory predictions, given the physical changes that occur during maturation such as lengthening of the vocal tract, descent of the larynx, and increased lung capacity. Considering the species which did not follow these general developmental trends, many belong to the same taxonomic groups. These differences will be discussed and evaluated in an attempt to explain the observed changes.

Cetacea and Chiroptera

Echolocation appears to develop similarly in cetaceans and bats. From birth, dolphins produce narrowband whistles and broadband burst pulse sounds, while new born bats produce low frequency FM isolation calls. With age, echolocation clicks appear and overtime the clicks increase in frequency and decrease in duration. These click trains seem to increase in duration and number of clicks, and decrease in repetition rate with age. Cetacean clicks

stabilize and resemble adult echolocation within the first year, potentially within a few months. Bat pups produce adult echolocation clicks as early as one month, typically corresponding with the bats first few flights.

While in most mammals call frequency decreases with age, in bats and dolphins frequency increases. This may be attributed to their unique use of echolocation. In order to gain enough resolution for tasks such as prey detection and navigation, the animals need to produce directional, high frequency clicks (Tyack, 1998). While young animals may not be physically capable of producing such high frequencies, the physical structures for sound production may develop throughout maturation, allowing for higher frequency production with age and enabling better resolution, adult echolocation.

The developmental similarities between cetacean and bat echolocation is particularly interesting given the ability convergently evolved using different sound production mechanisms. Bats produce calls similarly to most other mammals using vibrations in the larynx, where the cricothyroid muscle contracts to tense the vocal folds and relax as the click is produced (Suthers & Fattu, 1973). Odontocetes have a unique sound production system where calls are generated by phonic lips within the rostrum through a system of air sacs and fat compartments (Cranford, 2000). Currently there are no known predictions for how echolocation clicks physically develop in cetaceans. As tension of the vocal folds/phonic lips presumably determine the clicks frequency, changes with age could be attributed to development of structures and muscles which control vibration rate (Dearolf et al., 2000; Killebrew et al., 2001). Similarly, maturation of the larynx and associated muscles may facilitate echolocation development in bats (Moss, 1988). Alternatively, in bats changes in call frequency correlate with high frequency hearing development (Rübsamen, 1987), and development of call production could thus be limited by auditory perception. This does not seem to be true from the few studies of young cetaceans, which appear to have upper frequency hearing limits comparable to adult animals (Nachtigall et al., 2005; Nachtigall et al., 2007).

While cetacean and bat click parameters change similarly with development, some differences are apparent. In cetaceans, calves continue to

produce burst pulses and whistles throughout their life, suggesting clicks are added into the repertoire as a separate call type. In bats, isolation calls appear to evolve into clicks by progressively changing in frequency and duration, and pups do not continue to produce isolation calls after echolocation has developed. Additionally, in cetaceans echolocation appears to develop at a slower rate than in bats. Some observations have seen adult-like echolocation within seven weeks (Favaro et al., 2013) while in others clicks fail to appear for up to five months (Manoukian et al., 2002). Dolphin calves have extended periods of developmental care, up to two years, which may support longer periods for echolocation to fully develop (Reynolds et al., 2000). In contrast, bats develop adult echolocation as early as three weeks (Moss, 1988). Bats may develop more quickly due to shorter maternal care periods; in some species, bats wean as early as within one week (Buchler, 1980).

While developmental studies have focused on echolocation in bats, in cetaceans whistle changes have also been investigated. Overall, whistles appear to increase in amplitude, frequency, frequency modulation, number of loops, and duration with age. These whistles appear to be generated by vibrations of the phonic lips (Madsen et al., 2012) and given their unique sound production physiology there are no known theories for how call development is predicted to develop in these species.

Mysticetes are thought to produce calls comparably to terrestrial mammals, where sound is generated by air moving through U-folds at the end of the laryngeal sac (Reidenberg & Laitman, 2007). Thus, mysticetes might be expected to follow predictions of source-filter theory (Mercado et al., 2010) and have similar developmental patterns to other mammals. However, to date no studies have examined vocal development in any mysticete and these predictions cannot be confirmed.

Further studies on cetacean call ontogeny are difficult given the unreliable access to animals in the ocean environment. However, current advancements in recording and observation technologies are providing further opportunities to examine normal call development in wild species. Future research should continue to monitor wild populations, utilize opportunistic recordings of young

stranded animals, and where possible examine physiological changes in sound production structures through diagnostic imaging and dissection techniques to gain a better understanding of their unique sound production mechanisms and development.

Pinnipeds

Varying developmental trends have been observed in pinniped calls. Frequency parameters increased in Mediterranean monk seals, bearded seals, and sub-antarctic fur seals, decreased in Southern elephant seals, Hawaiian monk seals, and harbour seals, and in leopard seals there were no significant changes with age.

Currently there are no known explanations for the difference in developmental trends among pinnipeds. One possible explanation may be the varying developmental periods and growth rates between pinniped species. Pinnipeds have varying periods of maternal care; some nurse pups for as little as four days, while others have maternal care for up to three years (Bowen, 1991). In species with short maternal care periods, the young typically gain large amounts of weight in very little time. For example, in hooded seals (*Cystophora cristata*), pups double in weight over four days (Bowen et al., 1987). In species with longer maternal care periods, pup growth appears to be slower (Schulz & Bowen, 2004). After pups wean from their mothers, most species go through a fasting period in which pups lose weight before beginning to forage independently. If weight correlates with call frequency, then different call trends would occur across species during different growth periods.

The majority of studies only looked at a brief period of development just after birth, when pups were either nursing or just weaned. Although in many species pups grow rapidly during this short nursing period, most growth is attributed to blubber gain rather than physical development; for example in the hooded seal 70% of their weight gain is attributed to fat (Bowen et al., 1987). In other mammals the negative correlation between frequency and weight is predominantly attributed to changes in vocal tract length (Fitch, 1997). Given their variable physiological development, frequency changes might be better

predicted by changes in body length as opposed to weight measures of size in seal pups.

Developmental changes in juveniles, as opposed to pups, may also be more consistent across seal species. Sanvito et al. (2008) looked at the development of agonistic vocalisations in male Southern elephant seals, from juveniles through adulthood. As the juveniles aged, formants had higher intensity towards lower frequencies, as would be expected due to physical growth. These frequency changes occurred through the first seven to eight years, at which point the animals reached physical maturity (Sanvito et al., 2008). Interestingly, male leopard seals were also examined from juveniles through adulthood, but did not show any developmental frequency trends (Rogers, 2007). This difference between elephant and leopard seals may be attributed to size; upon adulthood, male Southern elephant seals weigh up to 3,500 kg, while male leopard seals only reach up to 400 kg (Weckerly, 1998). The drastic change in body size from juvenile to adults may explain the more prominent call difference in Southern elephant seals, while the smaller leopard seals experience less growth and may thus have more consistent calls.

Comparing vocal ontogeny in pinnipeds across studies is difficult as they examined different developmental periods. Some studies focused on neonatal calls when pups were still with their mothers (Job et al., 1995; van Opzeeland & van Parijs, 2004; Khan et al., 2006; Muñoz et al., 2011), some extend into the weaning period (Charrier et al., 2003) while others examine juveniles into adulthood (Davies et al., 2006; Rogers, 2007; Sanvito et al., 2008). This presents additional difficulty when trying to examine developmental trends across species, especially when the different species have such different growth periods. Given the variable trends observed across pinnipeds, more studies examining call development in conjunction with morphological growth from birth through adult development would be valuable. Currently only one study has examined an otariid species, and there is no information available concerning the other pinniped family: walruses. Further studies comparing development across all three pinniped families would be useful.

2.3.6 Additional Functions and Limitations

Previously I discussed some of the developmental trends that occur across mammalian species, and posited possible explanations for the similarities or dissimilarities in particular groups. In this section I will examine additional functions of vocal development. Evolutionary pressures, such as predation, may have driven selection away from changes predicted by source-filter theory and explain some of the observed differences. 'Babbling' also appears in many species during repertoire development. Such variability in infant calls may be partially attributed to physical maturation and the limitations of immature vocal production systems, however learning through self and social feedback may also contribute to development of adult repertoires. I will then discuss the limitations of the studies I have presented; many of the studies reviewed here have had limitations which should be kept in mind when interpreting trends.

Evolutionary Adaptation

In many mammals, calls appear to be honest indicators of animal size (Fitch & Hauser, 2003). Acoustic features can provide information about the caller, and in several species individuals with lower fundamental and/or formant frequencies have increased reproductive success and dominant social standing during aggressive interactions with rivals (e.g. North American bison, *Bison bison:* Wyman et al., 2012; domestic dogs: Taylor et al., 2010; koalas, *Phascolarctos cinereus*: Charlton et al., 2012 & 2013; Australian sea lions: Charrier et al., 2011). The use of such cues can be difficult to evaluate, given the large amount of information that can be encoded in different acoustic parameters. For example, in the red deer lower formant frequencies are thought to be used to assess rivals and to select preferred mates (Reby & McComb, 2003; Reby et al., 2005; Charlton et al., 2007a & b). However, deer hinds showed no preference to playbacks of male calls with varying fundamental frequencies (Charlton et al., 2008) except during oestrus, when females appear to prefer males with higher fundamental frequencies (Reby et al., 2010). While higher fundamental frequencies may be used during mate selection, playback

experiments show no difference in response by other males, suggesting that the fundamental frequency is not used as a cue when assessing rival males (Garcia et al., 2012).

Alternatively, some systems may favour dishonest signalling. In ground squirrels, juveniles are vulnerable to predators that do not pursue adults. If juvenile calls differ from those of adults, and are thus honest signals of age, predators could eavesdrop on calls to locate prey. Juvenile squirrels that mimicked the calls of adults could then be favoured by reduced predation. This is consistent with observations of alarm calls in ground squirrels, which have no significant changes in call parameters with age or size (Matrosova et al., 2007). Such pressures may favour young animals with particular call features. Future research examining call development should also consider life history and environmental pressures that may favour varying vocal signals.

Babbling

During development, human infants go through a 'babbling' stage during which their vocalisations resemble the consonant-vowel subunits of adult speech, but differ in that they are produced in rhythmic, repetitive bouts and are highly variable in their structure and the contexts they are made (e.g. Oller et al., 1976; Locke, 1986; Oller & Eilers, 1988). Regardless of culture, babbling occurs universally in humans (Oller, 1980). As previously discussed, a comparable developmental phase occurs in songbirds, where infants produce 'subsong', which lacks the typical structure of adult calls, and upon maturation begin to produce 'plastic' song that is more similar to adult repertoires but are overproduced (e.g. Marler & Peters, 1982; Marler & Nelson, 1992; Catchpole & Slater, 1995). With practice and feedback from conspecifics (typically in the form of 'duets') 'plastic' songs later crystalize into adult song (e.g. Marler & Peters, 1982; Marler & Nelson, 1992; Catchpole & Slater, 1995). At this time, specific songs may be selectively socially reinforced, leading to certain songs being reproduced as adults while others are discarded through 'action-based learning' (Marler & Nelson, 1992).

Recent studies have investigated the occurrence of babbling in several mammalian species including primates (e.g. Hauser, 1989; Elowson et al., 1998a & b; Snowdon & Elowson, 2001), bats (Knörnschild et al., 2006) and dolphins (McCowan & Reiss, 1995). The most thorough of these investigations has been of the pygmy marmoset (*Cebuella pygmae*), which has a short 'babbling' phase which occurs in infants but disappears upon weaning at approximately 20 weeks of age (Elowson et al., 1998a & b). Pygmy marmoset 'babbling' bouts were found to be long duration series of variable calls consisting of several different types (frequently duplicated), produced without apparent context (Elowson et al., 1998a & b).

Babbling is primarily thought to facilitate individuals matching their own vocal output to an auditory template through self-feedback (e.g. Knörnschild et al., 2010). Babbling may also facilitate learning; for example young ground squirrels frequently produce predator alarm calls that have a similar structure to adult calls, but produce them during inappropriate contexts (i.e. they do not exclusively produce the calls in the presence of a predator) (Mateo 1996a & b). Babbling may allow for young animals to learn to appropriately produce their calls through experience.

To a degree, babbling may also be influenced by physical development, such that immature vocal production systems are incapable of producing species specific calls, but with maturation more finite control over calls' develops. Infants who 'babble' more frequently are also thought to receive more attention from care-givers (Elowson et al., 1998a & b); however this does not appear to be true in all species, such as bats (Knörnschild, 2006).

To date few studies have investigated the structure and usage of infant calls through development into adult repertoires. Given the diverse species which show a 'babbling' stage, it may be a prevalent precursor to development of adult repertoires across mammals, especially in species with complex vocal repertoires. Future research should consider salient measures indicative of 'babbling' stages, such as variability, rates and usage when investigating call development.

Methodological Limitations

Comparing developmental trends within and across species is difficult given the methodological differences across studies. Here I discuss how such differences make interpretation and comparison of developmental trends difficult.

Recent advancements in sound recording and analysis technology have made objective and sensitive acoustic measurement widely available. Before these technologies were accessible, many developmental studies were limited to examination of simple parameters such as call rates and durations, or relied on human classification of call types to examine how repertoires change through development. Such methods are limited; early studies on non-human primate vocal development found that infant calls appeared to be fixed at birth, with very little change through maturation (Hauser, 1989). However, using current technology to more sensitively examine acoustic parameters has shown several developmental trends. Seyfarth & Cheney (1997) found that of the past literature examining call changes in primates, 79% of studies published before 1987 found call repertoires were fixed from birth. In comparison, after 1987 71% of studies found that calls changed with age. Such differences in technology inhibit comparison across studies.

Even among recent studies, comparing results can be difficult because many do not measure the same acoustic parameters. Of the studies reviewed here, 23% (16) did not directly state which acoustic parameters they measured (though some could be inferred from the results and discussion). Studies which did explicitly state the acoustic parameters they measured varied widely; for example, when looking at frequency parameters some examined fundamental, peak and/or formant frequencies. Some measured these parameters at different locations across the calls, such as at the start, middle or end, while some examined the maximum, minimum and/or mean values from taken from the entire call.

Future research examining call development would benefit from standardization of clearly defined acoustic parameters. Chosen parameters should correspond as much as possible to previous literature to aid comparison,

but should also consider the biological relevance of calls. For example, in dolphins and birds, frequency modulation patterns are behaviourally significant and should be measured. In most mammals, fundamental and formant frequencies appear to be honest signals of age, size and sex following source-filter theory predictions. Thus, future studies in most mammals should particularly consider how energy is distributed across call frequencies, where peaks in the spectrum occur, and perhaps the size of frequency bandwidths.

The studies reviewed here have also recorded animals in different behavioural contexts. For instance, many studies recorded infants when isolated and/or reunited with their mothers (e.g. Moss et al., 1997; Hammerschmidt et al., 2000). Other studies elicited calls by physically handling or approaching the animals (e.g. Colvin, 1973; Shipley et al., 1988; Sanvito et al., 2008). Others passively recorded individuals when in normal living conditions, either in captivity (e.g. Darden & Dabelsteen, 2006; Favaro et al., 2013) or by following wild animals (e.g. Job et al., 1995; Knörnschild et al., 2010). Presumably, the animals recorded in different contexts would also have varying behavioural states, which could affect call production. Stress, such as might be expected from separation of infants from their mother or exposure to humans, is known to alter many call parameters such as rate, duration, amplitude, frequency and modulation (e.g. Briefer, 2012; Briefer et al., 2015). However, with repeated exposures and experience, the stress of these events may be reduced and lead to call changes that could be misattributed to development. While it can be difficult to control motivational states of recorded animals, effort should be taken to document behaviour in conjunction with calls to help facilitate comparison between studies. When possible, passive recordings may be ideal when examining call development as changes attributed to motivational states induced by the presence of humans could alter results.

Similarly, development of calls may differ between captive and wild populations. In captive environments, individuals are presumably kept in smaller spaces and in closer proximity than in wild populations. Such differences may alter call production as the behavioural relevance of producing social calls in

such close proximity may not be comparable. Similarly, one might expect in echolocating species that use clicks for navigation, captive animals may have reduced click rates as visibility may be more reliable than in natural environments.

This survey was restricted to examining developmental trends of objectively measured acoustic parameters. Many studies were excluded from this review as they only discussed developmental changes in terms of the appearance or disappearance of specific call types. As previously discussed when examining grey seal repertoires, human classification of call types may be inconsistent and inhibit comparison between studies. However, there are benefits to the method. Many of the studies examined objectively measured changes by specific call types (see Table 2.1). Such type based studies classify calls using biologically salient categories which appear to be behaviourally significant to the animals. For example, in dolphins the separation of calls into clicks, burst pulses and whistles (particularly signature whistles) is behaviourally relevant to the species and important to consider when evaluating call ontogeny. Developmental trends might be expected to vary across such call types. Examining all calls produced by an animal may mask relevant developmental changes, when such trends may be apparent within call types. Future studies examining call development would benefit by taking into account biologically relevant sound categorization, but should also use objective classification and labelling methods to facilitate consistency across studies.

2.4 Summary

Here I have reviewed the current literature concerning call development, considering physical maturation and learning, in mammals. As illustrated by 'babbling', call development is a complex process by which calls evolve to adult repertoires due to both physical maturation and learning. Throughout the rest of this thesis, I will examine the vocal development of the grey seal. By examining call changes that occur across seals, trends that may be attributed to physical development and morphology are identified. Then, by examining call changes that occur within individual animals, involving when and how sounds are

produced or used, changes that can be attributed to learning are identified. Considering how physical development and learning both play a role in call development will provide a more comprehensive understanding of how the complex vocal repertoires of pinnipeds arise.

Chapter 3
Vocal Development in the Grey Seal

3.1   Summary

Here, I examine developmental call changes in the grey seal that can be attributed to physiological maturation. 93 grey seals, ranging from pups to adults, were recorded on the Isle of May. Three of these pups were recorded regularly from birth through their first year of life. The seals' sounds were compared across age and sex groups. In general, calls decreased in frequency and increased in duration with age, and call frequency was higher and duration shorter in females than in males. Additionally, the total bandwidth of calls increased with age, and noise and variability decreased with age. This developmental pattern is similar to what has been observed in most other mammals, and follows the predictions of source-filter theory.

3.2   Introduction

In chapter 2, I reviewed the current literature on vocal development in mammals. Generally, most mammals follow the developmental predictions of source-filter theory (e.g. Fant, 1960; Titze, 1994) such that maturation correlates negatively with frequency and positively with duration. This is attributed to physiological changes such as the lengthening of the vocal tract, descent of the larynx, and increased lung capacity. However, not all mammalian species followed these developmental trends, such as the pinnipeds. With age, frequency increased in Mediterranean monk seals (*Monachus monachus)*, bearded seals (*Erignathus barbatus)*, and sub-antarctic fur seals (Muñoz et al., 2011; Davies et al., 2006; Charrier et al., 2003) but decreased in Southern elephant seals (*Mirounga leonina*), Hawaiian monk seals (*Monachus schauinslandi)*, and harbour seals (*Phoca vitulina*) (Sanvito et al., 2008; Job et al., 1995; Khan et al., 2006). However, in leopard seals (*Hyrdrurga leptonyx*) there appeared to be no significant age related changes in frequency variables (Rogers, 2007).

Potentially, these developmental differences between pinniped species may be attributed to morphological growth; pinnipeds have variable periods of

maternal care and growth rates. Pups with short maternal care periods typically gain large amounts of weight in very little time, while species with longer maternal care periods have slower growth (Schulz & Bowen, 2004). If weight correlates with call frequency, then different call trends may occur across species during different growth periods. However, rapid growth in some species can be attributed to an increase in body fat, and may not correspond with growth of the vocal tract. Thus, mass based predictions for pinniped vocal development are difficult to anticipate. Other measures, such as length, may be more reliable indicators.

Grey seals are amongst the most vocal of pinniped species, producing varied aerial and underwater calls. While the vocal repertoire of this species has been examined (Asselin et al., 1993; McCulloch, 1999), little is known about call development. Grey seal pups have been reported to produce only one call type; a periodic harmonic sound typically referred to as a 'mother attraction' call (Asselin et al., 1993; Caudron et al., 1998; McCulloch, 1999; McCulloch et al., 1999). Given the complexity of the adult repertoires more information on how calls develop from birth is needed.

Here, I documented the vocal development of grey seals. If grey seals follow source-filter theory predictions, their calls are anticipated to decrease in frequency and increase in duration with age. Grey seals are a sexually dimorphic species, with males being larger than females. Males are thus expected to have lower frequency parameters and longer call durations than females.

3.3   Methods
3.3.1   Subjects

Vocalisations of grey seal pups were recorded from birth through the first year of life. Between birth and weaning, recordings were made of a wild population on the Isle of May (Firth of Forth, Scotland). After weaning some of the pups were taken to the Gatty Marine Laboratory (St. Andrews, Scotland), a licensed captive research facility. These pups remained in captivity for 12 months before being released back into the wild.

On the Isle of May, six pups (three male and three female) were opportunistically selected as focal animals based on location and birth date. Only animals near the periphery of the colony were recorded to minimize disturbance to the seals. Researchers observed all of the focal pups being born or just after birth (as indicated by blood on the pup's fur and the presence of the placenta nearby). Grey seal pups are typically weaned between two and three weeks of age (Mellish et al., 1999) and researchers left the Isle at the beginning of December due to time and weather restrictions. Consequently, only pups born between November 4th and 12th were selected because they were expected to wean just before the researchers left.

Of the six pups selected, two (both female) weaned earlier and one (male) weaned later than anticipated. These three pups were not taken into captivity. The three pups who weaned within the expected time period were tagged and transported to the Gatty Marine Laboratory ("Clark", tag numbers 73094/5, "Bud", tag numbers 73010/1, and "Zola", tag numbers 73254/5) on December 6th, 2011. Two additional pups were also brought to the facility ("Yoza", tag numbers 72839/40, and "Angel", tag numbers 72824/5).

3.3.2   Recording: Isle of May

A total of 175 hours of recording were made on the Isle of May (mean time each focal pup was recorded = 28.2 hours, ± 8.7). Daily in-air acoustic and video recordings were made of each focal pup from birth through December 3rd, 2011 using a Sennheiser MKH 416 P48 directional microphone (frequency response 40 Hz to 20 kHz, sensitivity at 1 kHz 25 mV/Pa +/-1 dB). Recordings were high pass filtered at 100 Hz to exclude ambient wind noise and digitized into .wav format using a Marantz Pro Solid-state recorder PMD671 (sampling rate 96 kHz, 24 bit). Concurrent video recordings were made using a Sony DCR-HC96E digital video camera. In the few occasions that the Marantz recorder failed, acoustic recordings were taken from the video recorder for analysis (sampling rate 48 kHz, 16 bit).

The microphone and video recorder were placed as close to the focal animal as possible without visibly disturbing the seals. The focal animal was

never more than 10 m away from the recording equipment. In cases where the animal was near a rock wall, the microphone was placed as near as 0.5 m from the focal pup by setting up all equipment from behind the wall, hidden from view. A researcher observed all recording sessions from either behind a rock wall or a distance of at least 10 m away from the focal seal. Each pup was recorded between 30 and 120 minutes each day, depending on weather. Recordings were not made in adverse conditions (rain and/or wind over 25 mph). Recordings were limited to sunlit hours (approximately 0700 to 1700) to minimize disturbance to the colony and for the researcher's safety. This limitation was not expected to be problematic as previous investigations have found that overall grey seals do not vary their vocal behaviour with time of day or lighting conditions (Asselin et al., 1993).

During each recording session, the identity and relative location of all animals within 30 m of the focal pup were noted. Adult females were identified by unique coat patterns (Redman, 2002; Paterson et al., 2013) and pups were identified through: 1) association with their mother, 2) approximate developmental stage and 3) location. Adult males were transient and could not be individually identified. Thus, only recordings from the same session could be attributed to a single male. While recordings focused on the six pups selected, vocalisations of the pups' mothers and neighbouring animals were concurrently obtained.

Additional opportunistic recordings were made of Zola, Yoza, and Angel during a separate study (Robinson, 2014) on the Isle of May. During this study individual pups were identified with flipper tags and unique paint markings applied to the pup's back. Two groups of ten pups (including Zola, Yoza, and Angel) of approximately the same age were captured and placed into temporary holding pens. The two groups differed temporally (Yoza took part in the first group, tested from November 19th through 26th, and Zola and Angel took part in the second group, tested from November 27th through December 4th). Over the course of eight days, every other day each pup was captured, intravenously injected with either saline or oxytocin, and placed into a testing enclosure with another pup. The pups were observed for one hour before being recaptured and

replaced into their respective holding pens. While in the testing enclosure, behaviour was monitored using video recordings made with a Panasonic HDC-TM60 digital video camera. Acoustic recordings were concurrently made using the Sennheiser MKH 416 P48 microphone and Marantz Pro Solid-state recorder PMD671. In total, each pup was placed in the testing pen four times (Robinson, 2014). At the end of the trials, all seals were released and those brought into captivity were re-captured as free roaming animals.

3.3.3 Recording: Gatty Marine Laboratory

Upon arrival at the Gatty Marine Laboratory, recordings continued using the same equipment with the addition of underwater monitoring using a HTI 96-min hydrophone (frequency response 2 Hz to 30 kHz, sensitivity -201 dB re: 1V/µPa). At minimum, one hour of simultaneous in-air and underwater recordings occurred weekly during daylight hours, with additional recordings taken opportunistically. Recordings were not made during adverse weather conditions (rain and/or winds over 25 mph) or during periodic pool maintenance. When in captivity, individual seals were visually identified with flipper tags.

Seals were housed variably in three enclosures including a large rectangular pool (42 x 6 x 2.5 m) and two circular pools (3 x 5 x 2 m). Each pool was attached to dry haul out areas and separated from each other by metal fencing. All of the seals were moved regularly between these pools via a series of gates. The main pool connected to the small pool via an underwater gate, and the small pool connected to the isolation pool via a terrestrial gate. Recordings were made primarily in the isolation pool because 1) underwater sounds were assumed to be isolated to that pool as there was no underwater gate connecting it to the other pools and 2) an underwater housing was permanently affixed to the side of the pool and was used to protect equipment from the animals.

To evaluate how representative one hour samples taken during the day were of the seals' overall vocal behaviour, three 24 hour recordings were made using the same equipment (22 kHz sampling rate, 16 bit). It was not anticipated that any differences would be observed between the day and night samples as

overall grey seals have not been found to vary their vocal behaviour with time of day or lighting conditions (Asselin et al., 1993). While Asselin found that for most call types there were no diurnal patterns, the seals produced one underwater call type ('clicks') more often at night (Asselin et al., 1993). It was also possible that the presence of human researchers at the facility during the day may have affected vocalisation production.

Two of the captive seals (Clark and Zola) also participated in a separate study investigating vocal usage and production learning. Both seals were conditioned using positive reinforcement to target, station, and vocalise on cue. These vocalisations were shaped using operant conditioning to imitate novel stimuli and vary in acoustic variables such as number of calls, duration and frequency. Recordings were made of all training and test sessions. Further information regarding this experiment can be found in chapters 4 and 5. As training would be expected to alter the seals' call production, none of the recordings made during the vocal learning studies were included for analysis in this study. However, training may have inadvertently impacted the calls produced outside of these sessions.

3.3.4   Sound Analysis

Audio recordings were evaluated using Adobe Audition 2.0 (FFT size: 2048, frequency resolution: 46.87 Hz, time resolution: 10.66 ms, weighting function: Hamming, window width: 100%). The seal's vocalisations were compared to the video feed, and sounds corresponding to visible movements of the face, head or diaphragm of an identifiable animal were isolated for further analysis. Each continuous seal sound (without a break of more than 5 ms) at least 10 dB above background noise was counted as one call.

For each sound, background information was noted including: identity of the vocalizing animal, location, time, date, sex, and age. On the Isle of May, focal pups were spread across three different locations, separated by rock walls, referred to as: kirkhaven, tennis court, and cross park (Figure 3.1). If known, age of the animal in days was recorded. If not, age was estimated based on physical characteristics. Kovacs and Lavigne's (1986) five point scale

Figure 3.1. Map of the Isle of May showing study sites. Seals were recorded from three locations in the south of the island: A) 'kirkhaven', B) 'tennis court' and C) 'cross park'. Solid white lines indicate trial paths, while tan lines indicate rock walls which separated the study areas. Map taken and adapted from Scotland's National Nature Reserves official website (http://www.snh.org.uk/pdfs/publications/designatedareas/islemay%20visitor%20map.pdf).

Table 3.1. Seal pup age classification with characteristic description, mean age in days, and standard deviation (taken from Kovacs & Lavigne, 1986).

Age-Class	Description	Age in Days (Mean +/-SD)
I	Yellow tint to pelage; lacking coordination; neck, hips and ribs clearly visible; umbilicus present	2.4 +/- 4.4
II	Pelage white; improved coordination, shoulder to hip region filled out; ribs covered by a layer of blubber; umbilicus not present	4.8 +/- 3.1
III	Pelage white to light grey; fat sheath extends posteriorly from neck; body barrel shaped; lanugo intact except for slight loss in facial region toward the end of stage	12.1 +/- 2.9
IV	Lanugo being shed, exposing the juvenile pelage	16.0 +/- 3.0
V	Moulted pup, less than ~5% of the body surface still retaining lanugo	>21

was used to categorize pup age (Table 3.1), and older animals were identified as weaned, yearling, or adult based on pelage and size. Age in days was then estimated for each group using the average provided by the five point scale (e.g. stage 1 pups were averagely 2.4 days old). Adults were assumed to be the average age for sexual maturity (females: 1825 days, males: 1095, Hammill & Gosselin, 1995).

Previously, grey seal pups were described as having one call type: a periodic, harmonic call generally referred to as a 'mother attraction' call (Asselin et al., 1993; Caudron et al., 1998; McCulloch, 1999; McCulloch et al., 1999). However, many previous descriptions of vocal development in other species discuss the appearance and development of distinctive call types (e.g. Coscia et al. 1991; Pistorio et al., 2006; Stoeger-Horwath et al. 2007). In this investigation, recorded calls were primarily classified into one of three categories for analysis; 'periodic', 'noisy' and 'mixed' calls. Calls consisting of periodic signals with energy concentrated in comparatively narrow frequency bands were classified as 'periodic'. These calls contained at least one emphasized frequency band, but could also have several (such as with harmonic calls). Calls without a clear periodic structure with energy distributed across a comparatively wide frequency bandwidth were classified as 'noisy'.

Any calls that contained both 'periodic' and 'noisy' elements were classified as 'mixed'. Examples of these call types are shown in Figure 3.2. Classification into these categories was chosen for several reasons. These call types have been observed in the repertoires of wild grey seal populations, although previous investigations have further split these into smaller call categories (Asselin et al., 1993; McCulloch, 1999). As discussed in chapter 2, classification of these smaller categories has been inconsistent across studies. For this reason, my analysis focused on the broader categories. Subjective observation also suggests the categories are behaviourally distinctive; 'noisy' calls appear to be associated with aggression, while 'periodic' sounds seem to occur in variable social situations including aggression, mother and pup interactions, and mating (Asselin et al., 1993; McCulloch, 1999). Grey seals can also generalize and classify calls into these types, which is further discussed in chapter 4. Lastly, categorizing the seal calls into these groups was pragmatic for analysis as not all parameters could be measured for all call types. For example, as discussed in chapter 2, the fundamental frequency appears to change with development and may carry socially relevant information. While this parameter can be measured in 'periodic' calls, it cannot be measured in 'noisy' calls.

In addition to the analysis with these three call types, another call type was identified for separate analysis. The underwater repertoire of grey seals includes sequences of repetitive, multiple element calls such as 'rups', 'rupes', and 'trrots' (Asselin et al., 1993; McCulloch, 1999). Although McCulloch (1999) did not define how calls were classified as single versus multiple element calls, Asselin et al. (1993) defined these vocalisation sequences as calls emitted without a pause of more than two seconds. These previous investigations have described the 'rupe' as an underwater, repetitive, multiple element call (Asselin et al., 1993; McCulloch, 1999). One unit of the call consists of a broadband pulse and appears to be very similar to how the 'rup' is described. While McCulloch (1999) did not describe the difference between the 'rup' and this unit of the 'rupe', Asselin et al. (1993) noted a more harmonic structure in the 'rupe'. However, examples of 'rups' and the first unit of the 'rupes' provided by
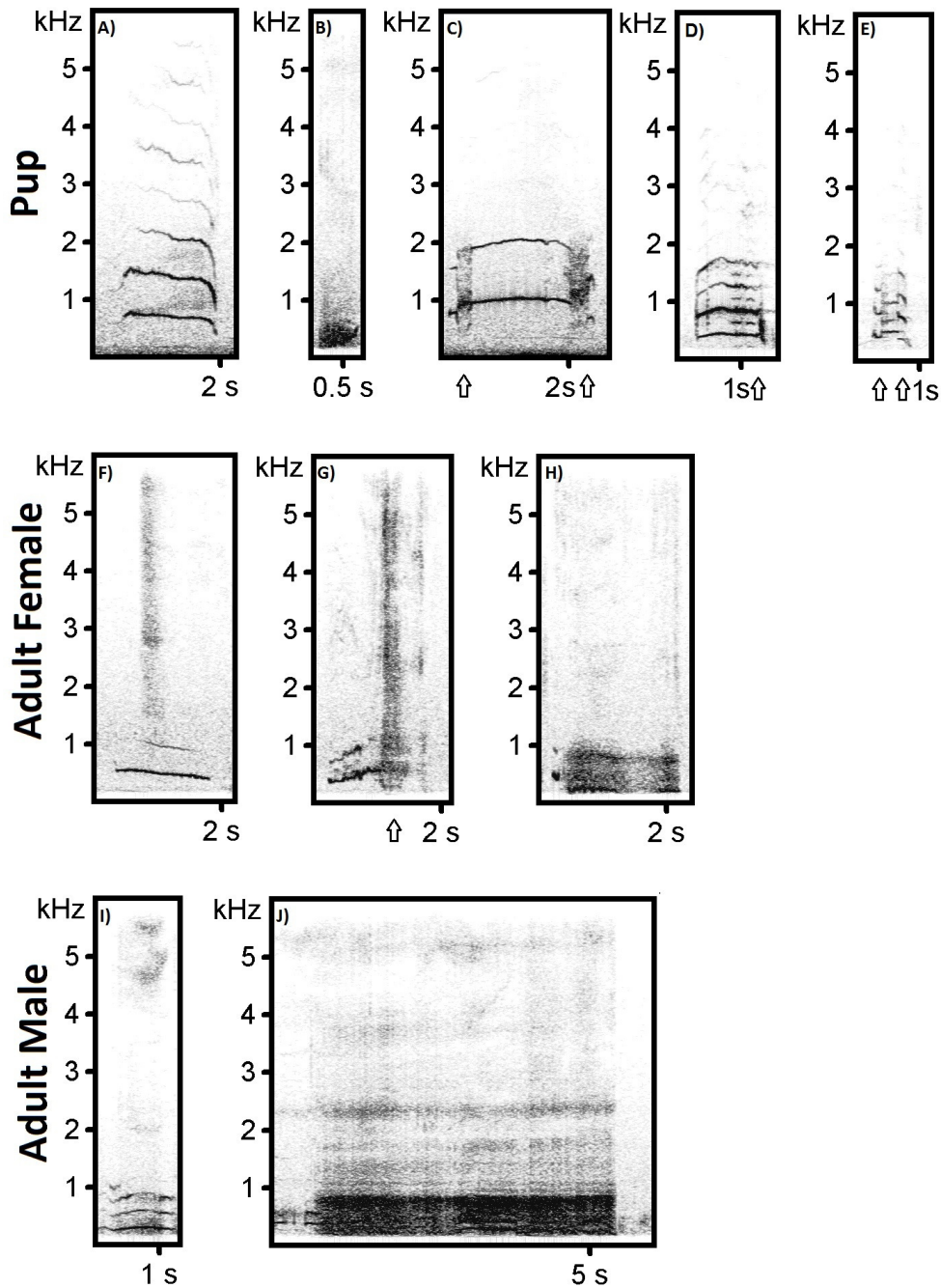
Figure 3.2. Spectrographic examples of grey seal pup, adult female and adult male calls. Time is shown in seconds on the x-axis, and frequency in kilohertz on the y-axis. Calls were classified into types for analysis including 'noisy' (B, H, J), 'periodic' (A, D, E, F, I), and 'mixed' (C, G). Examples also show nonlinearities including "limit cycles" (A, C, D, E, F, G, I), "subharmonics" (D), "frequency jumps" (E), and "chaotic" (B, C, G, H, J). Locations where nonlinearities transition within a call are indicated with an arrow. Spectrograms created in Avisoft-SASlab Pro (FFT size: 2048, frequency resolution: 46.87 Hz, time resolution: 10.66 ms, weighting function: hamming, window width: 100%). Audio examples are included in the supplemental material, listed in Appendix I.

McCulloch (1999) appear to be very similar, suggesting that perhaps the 'rup' may be one unit of the 'rupe'. The second unit of the call consists of a more periodic, harmonic component that is typically longer in duration. Although not directly discussed in previous investigations, the order of these two units appears to be flexible, and varies in several parameters such as unit duration, time between individual units and call sequences, frequency structure (such as an upsweep or downsweep), and number of unit repetitions within a sequence (Asselin et al., 1993).

Given that very little in known about these sequenced calls, additional analysis was used to further describe these specific calls by separating sequences (any calls produce consecutively without a break of more than two seconds) into individual units (any sound without a break greater than 5 ms) and subjectively classifying these units as being either a 'noisy' pulse or 'periodic' harmonic element. The 'noisy' sounds appear to be similar to previous descriptions of the 'rup' and one unit of the 'rupe', and the 'periodic' to those of the other unit of the 'rupe' (Figure 3.3). This sequence analysis occurred in addition to analysing the calls as a single 'noisy' call for the developmental analysis.

Frequency and time variables were measured using Avisoft-Saslab Pro 5.02.04 sonogram software. The same variables used by McCulloch (1999) and Asselin et al. (1993) were measured to facilitate comparison between studies, including duration, fundamental frequency, beginning frequency, end frequency, maximum frequency, and minimum frequency. Additional variables thought relevant to the description of the vocalisations were also included, such as signal to noise ratio (SNR), Wiener entropy, harmonic to noise ratio (HNR), peak frequency and nonlinearities (see table 3.2 for definitions). The SNR, which measured the call's energy in comparison to background noise, was used to evaluate the quality of recordings when selecting which calls were included for analysis. Any call with a SNR below two was excluded from analysis.

Previous investigations have typically measured acoustic parameters at set points across the duration of the call (most often the centre) or averaged measurements across the entire call. Here, I used both approaches. First, I
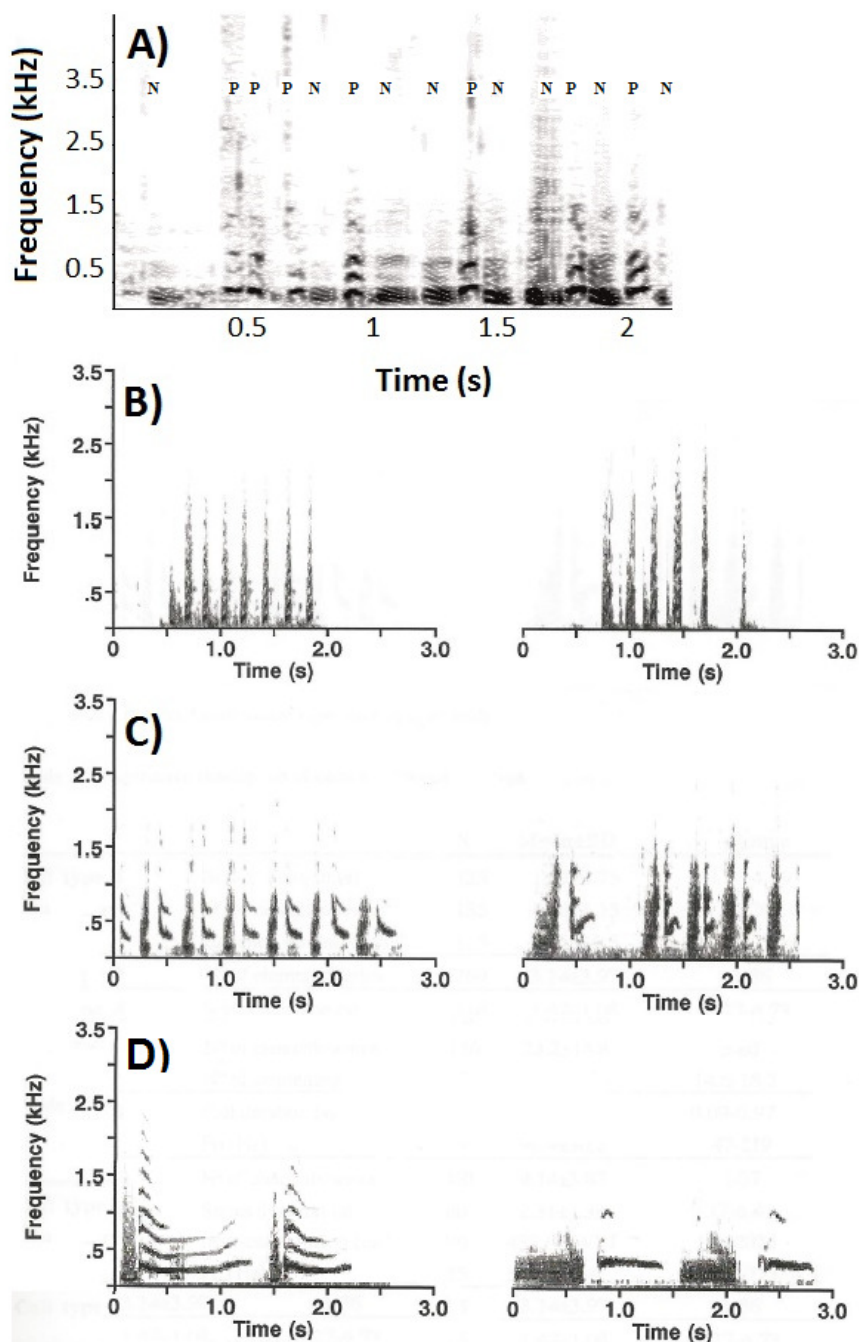
Figure 3.3. Spectrographic examples of grey seal underwater 'sequence' calls. Sequences show: A) 'noisy' (**N**) and 'periodic' (**P**) units, B) 'rups', C and D) 'rupes'. 'Rups' are similar to the noisy unit of the 'rupes', if 'rups' are defined as being variable in duration and frequency bandwidth. Spectrogram A was created in Avisoft-SASlab Pro (FFT size: 2048, frequency resolution: 46.87 Hz, time resolution: 10.66 ms, weighting function: hamming, window width: 100%). Audio example of recording A is included in Appendix II. Figures B), C) and D) were reproduced from McCulloch (1999).

Table 3.2. Definitions for the acoustic variables measured. Time is measured in ms and frequency in Hz. Parameters (indicated by a *) were measured at multiple points across the call (as discussed in the methods). Not all parameters were measured for all call types. For 'noisy' and 'mixed calls, 84 parameters were measured. For 'periodic' calls, 93 were measured.

Variable	Definition
# Frequency Spectrum Peaks*	The number of peaks in the frequency spectrum within 20 dB of the maximum peak.
First Frequency Spectrum Peak*	The frequency with the maximum amplitude within the first spectrum peak referred to in "# frequency spectrum peaks".
Peak Frequency*	The frequency with the highest amplitude.
Peak Frequency Bandwidth	Frequency difference between the maximum and minimum of the frequency spectrum peak with the highest amplitude (see above) within 20 dB of the spectrum's maximum peak.
Total Bandwidth*	Difference between the minimum and maximum frequency of the entire spectrum within 20 dB of the maximum spectrum peak
Maximum Frequency*	The highest frequency of the entire spectrum within 20 dB of the maximum spectrum peak.
Minimum Frequency*	The lowest frequency of the entire spectrum within 20 dB of the maximum spectrum peak.
Duration	Time from the beginning to the end of the call, within 35 dB of the maximum spectrum peak.
Distance to Max	Time from the start (within 35 dB of the maximum spectrum peak) to the maximum amplitude of the call
Signal to Noise Ratio (SNR)	The sound's average energy (1 Volt^2*sec), divided by energy of background noise. If SNR was < 2, the call was excluded from analysis.
Wiener Entropy*	The ratio of the geometric mean to the arithmetic mean of the spectrum. This is closer to zero for pure-tone signals and one for random noise (Tchernichovski et al., 2000).
Fundamental Frequency*	Lowest integer multiple of frequency spectrum peaks in a periodic, harmonic call.
Harmonic to Noise Ratio*	The dB ratio between the energy of the waveform and the energy between waveforms. The ratio is averaged first for each frequency bin, then the frequency bins are averaged across the entire call.
Nonlinearities	Presence/absence of nonlinearities, identified through visual evaluation of the spectrogram and categorized by type (see Figure 3.2).
Nonlinear: Limit Cycle	The spectrum is composed of a fundamental frequency with harmonics that are integer multiplies of that frequency (Tokuda et al. 2002).
Nonlinear: Sub harmonics	The presence of frequency peaks that are integer fractions of the fundamental frequency (i.e. ½ multiples) (Tokuda et al. 2002).
Nonlinear: Frequency Jumps	A break in the Fundamental frequency in which the vibration rate increases or decreases abruptly and discontinuously (Tyson et al. 2007).
Nonlinear: Chaotic	A noisy segment with no particular harmonics in the spectrum (Tokuda et al. 2002).
% Nonlinear	Proportion of the call containing nonlinearities (i.e. the duration of the nonlinearities divided by the call's total duration).

measured the parameters at five points across the duration of the call (i.e. at the start, one quarter, middle, three quarters and end of the call). Second, I sampled all parameters every 5 ms across the duration of the call, and from these points measured the overall median, maximum, minimum, mean and standard deviation of these samples. I choose to use both methods to facilitate comparison with previous studies which have utilized varying measures, as well as to thoroughly describe the changes occurring within the calls. Not all variables could be measured for all three call types. For example, 'noisy' and portions of the 'mixed' calls do not have fundamental frequencies, so this measure was not included for these call types.

3.3.5   Statistical Analysis

Source-Filter Theory Variablesprincipal

All statistical testing was conducted using R version 2.15.3. To examine if calls changed according to source filter theory predictions, a general linear mixed effects model ('glmer') was conducted using the 'lme4' package (1.0-4) (Bates et al., 2013). Separate models were conducted for the separate call types: 'periodic', 'noisy' and 'mixed'. For 'periodic' calls, three models were fit to predict changes in the average fundamental and peak frequency, and the duration of calls. For 'noisy' and 'mixed' calls, only average peak frequency and duration were fitted because fundamental frequency was not measured for these call types. As the distribution of these response variables was not always normal, gamma distribution was used. In all of the models, individual seal was included as a random effect, with age, sex, and their interaction as covariates. Age was classified into three groups: 'pups' consisting of pups with their mother (stages 1-4), 'weaned' consisting of pups no longer with their mother but younger than one year (stage 5), and 'adults' consisting of any animal older than one year. Model contrasts between age and sex were examined using the 'lsmeans' function and package (1-10.01), and multiple testing was accounted for with Tukey's test statistic (Lenth, 2013).

Phylogenetic Principal Components Analysis

Biological sounds are complex stimuli, and additional variables to those predicted by source-filter theory were examined. As discussed previously, for 'noisy' and 'mixed' call types, 84 acoustic parameters were measured. For 'periodic' calls, 93 acoustic parameters were measured. These variables were condensed into a smaller number of parameters with a phylogenetic principal component analysis (PCA) using the function 'phyl.pca' in the package 'phytools' (Revell, 2012). As PCA groups' calls similarly to K-means cluster analysis (Ding & He, 2004), it was anticipated that the similarities within individual animal calls would affect clustering. Phylogenetic PCA was used as it allowed for calls to be nested by animal and accounted for repeated measures by individual. This identified the variables that accounted for the majority of variation occurring within the data, and condensed variables onto a smaller number of components. The number of components retained was determined using a scree plot and the amount of variability accounted for in the data.

Linear Mixed Effects Models

The retained components were then used as predictor variables in linear mixed effects models ('lmer') using the package 'lme4' (1.0-4) (Bates et al., 2013). To again account for repeated measures, all models included animal as a random effect, and when retained, date nested within animal. Date was nested within animal as it was assumed that there would be more variability in the calls of individuals across days than in days across individuals. In all models, animal and covariates of age (in days) and sex were retained. Additional covariates, including signal to noise ratio, presence/absence of nonlinearities, and interactions between covariates, were chosen using a stepwise model selection procedure. Optimal models were selected by first examining the possible combinations of random effects (animal and date) within covariates, and then all possible combinations of fixed effects (age, sex, signal to noise ratio and nonlinearities). Model selection was made using second-order Akaike information criterion ('AICc') in the package 'MuMIn' (1.9.13) (Barton, 2013). As the principal components were numerical with both positive and

negative values, a 'Gaussian' distribution had to be used. While the principal components were all normally distributed, numeric covariates (age and signal to noise ratio) were log transformed for normality before analysis. The final model assumptions were checked using diagnostic plots of the residuals. As the goal for this analysis was to explore trends within the data, and the models were made using different predictor variables and data sets, no correction for multiple testing was applied (Rothman, 1990; Bender & Lange, 2001).

Cosinor and Sequence Analysis

Due to practical restrictions, recordings were primarily made during normal working hours (09:00-17:00). It was assumed that calls produced during these hours would be representative of normal repertoires; however it is possible that the seals exhibited diurnal variation in calling behaviour. To examine if there were any diurnal patterns in the seal calls, a cosinor test, 'cosinor.lm', was conducted using the package 'cosinor' (1.1) (Sachs, 2014). Analysis was conducted on the three 24-hour recordings made of the captive juvenile seals to examine if call rate followed model predictions of a sine wave, assuming a 24 hour time period. Separate models were constructed for in-air and underwater calls, using the total number of calls produced and the number of 'noisy', 'periodic' and 'mixed' calls as predictors, with the additional call category 'sequence' calls for the underwater call analysis.

A separate analysis was conducted for these 'sequence' calls. Individual sounds within each sequence were subjectively classified as being either a 'noisy' pulse or 'periodic' harmonic element. The order of these call units within each sequence was analysed using sequence analysis 'seq' with the program 'TraMineR' (Gabadinho et al., 2011; Bürgin & Ritschard, 2012; Ritschard et al., 2013).

3.4   Results

In total 3,879 calls, produced by 93 seals, were analysed. See table 3.3 for the number of animals sampled at each age class by sex. Table 3.4 shows the average number of calls recorded per individual. Notably, the number of calls

Table 3.3 Number of animals sampled at each age by sex. * Note, these totals do not reflect the total number of animals recorded, but instead the number of animals sampled at each age. Age was classified using Kovac and Lavigne's (1986) 5 point scale to classify pup ages. Stage 5 pups are classified as weaners 'W' (21 to 50 days old). Older animals were classified as juveniles 'J' (51 days to 1 year old) or adults 'A' (older than 366 days old).

Age	Female	Male	Unknown	Total*
1	4	2	1	7
2	4	3	3	10
3	4	3	11	18
4	4	3	14	21
W	4	2	14	20
J	1	3	0	4
A	31	7	0	38
Total*	52	23	43	

Table 3.4 Mean (and standard deviation) for the number of calls recorded from each animal by age (pups < one year, adults > one year) and sex. Averages are shown for all animals combined in the top portion table, and for focal pups and their mothers in the bottom portion of the table. The number of calls recorded varied greatly between animals as a larger number of calls were recorded for focal animals than for neighbouring individuals.

ALL ANIMALS	Pups	Adults
Male	325 ± 255	16 ± 17
Female	220 ± 314	26 ± 43
Unknown	23 ± 29	N/A
FOCAL ANIMALS		
Male	425 ± 114	N/A
Female	397 ± 284	126 ± 33

recorded per individual varied greatly as more calls were recorded for the focal animals and their mothers in comparison to neighbouring individuals (Table 3.4). Table 3.5 lists the total number of calls analysed by age and call type.

Table 3.5. Total number of seal calls analysed by call type, sex, and age group. Pup age was classified using Kovac and Lavigne's (1986) scale. Stage 5 pups are classified as weaners 'W' (21 to 50 days old). Older animals were classified as juveniles 'J' (51 days to 1 year old) or adults 'Ad' (older than 1 year).

		Call Types			
	Age	Periodic	Noisy	Mixed	SubTotal
	1	149	26	7	182
	2	398	42	27	467
	3	59	25	28	112
Female	4	1	10	0	11
	W	0	36	0	36
	J	35	17	1	53
	Ad	499	216	129	844
	SubTotal	1141	372	192	1705
		Call Types			
	Age	Periodic	Noisy	Mixed	SubTotal
	1	130	37	5	172
	2	145	52	12	209
	3	135	35	26	196
Male	4	95	71	11	177
	W	19	113	13	145
	J	165	9	0	174
	Ad	21	54	10	85
	SubTotal	710	371	77	1158
		Call Types			
	Age	Periodic	Noisy	Mixed	SubTotal
	1	8	1	0	9
	2	72	5	8	85
	3	219	26	26	271
Unknown	4	392	75	50	517
	W	25	97	12	134
	J	0	0	0	0
	SubTotal	716	204	96	1016
	TOTAL	2567	947	365	3879

3.4.1　Source-Filter Theory Variables

Overall, the seal calls varied by age and sex as predicted by source filter theory (GLM models, Table 3.6, Figure 3.4). The fundamental frequency of 'periodic' calls decreased with age by ~ 133 Hz between nursing and adulthood. Males produced lower fundamental frequencies than females by ~ 95 Hz. However, there was no significant difference in fundamental frequency between nursing and weaned female pups (Figure 3.4). Peak frequency also decreased with age, but the rate varied by call type. For 'noisy' calls, peak frequency decreased by ~ 123 Hz from nursing to adulthood. However, there were no differences between sexes. For 'periodic' calls, peak frequency decreased with age at a faster rate, changing by ~ 257 Hz from nursing to adulthood. There was no difference in peak frequency by sex between nursing pups, but in weaned pups and adults peak frequency was lower in males. For 'mixed' calls, peak frequency decreased between nursing pups and adults by ~ 317 Hz. There was no difference between weaned pups and adults, or between sexes. However, this result should be interpreted with caution as there were fewer calls analysed for this model. For example, there was only one call from a weaned female pup. Such a small sample size may have resulted in low power to detect differences (table 3.5, Figure 3.4).

Call duration generally increased with age, and was longer in males than in females. For 'noisy' sounds, calls were ~ 0.75 seconds shorter in nursing pups compared to adults. Duration did not vary between sexes for nursing or weaned pups, but was longer in adult males than in adult females. For 'periodic' calls, nursing pup calls were ~ 0.7 seconds shorter than those of adults, but there was no difference between weaned and adult calls. Lastly, for 'mixed' calls, there was no difference in duration between nursing and weaned pups, but adult calls were ~ 1.0 second longer. There was no sex difference for duration of 'mixed' calls at any age, though this should be interpreted cautiously because sample size for some groups was small (table 3.5, Figure 3.4).

Table 3.6. Generalized linear mixed effects models (gamma distribution, inverse link) showing how average fundamental frequency, peak frequency and duration varied with age and sex by call type. Age was categorized into three categories; nursing pups 'P' (stages 1-4), weaned pups 'W' (stage 5 to 1 year) and adults (older than 1 year). Individual was the random effect. For age, adults were referential in the model. For sex, females were referential in the model. Significant (p < 0.05) variables are shown in bold.

Fund. (Periodic)

	Est.	CI 2.5%	CI 97.5%	P
Intercept	368.37	359.69	377.05	<2e-16
Age (P)	133.78	122.10	145.46	<2e-16
Age (W)	119.26	85.49	153.04	6.01E-12
Sex	-95.42	-140.57	-50.27	3.55E-05
Age(P): Sex	26.61	-20.01	73.22	0.263
Age(W): Sex	9.46	-48.04	66.97	0.747

Peak

		Noisy Est.	CI 2.5%	CI 97.5%	P		Periodic Est.	CI 2.5%	CI 97.5%	P		Mixed Est.	CI 2.5%	CI 97.5%	P
Peak	Intercept	486.61	458.09	515.13	<2e-16	Intercept	440.37	424.67	456.08	<2e-16	Intercept	417.74	380.15	455.32	<2e-16
	Age (P)	123.37	74.21	172.53	1.03E-06	Age (P)	257.28	236.14	278.41	<2e-16	Age (P)	317.83	251.81	383.86	<2e-16
	Age (W)	88.24	25.32	151.16	0.0061	Age (W)	173.91	112.80	235.02	2.74E-08	Age (W)	445.26	30.12	860.41	0.0357
	Sex	-55.26	-118.66	8.14	0.0874	Sex	-66.16	-147.85	15.52	0.1123	Sex	36.04	72.50	0.50	0.6196
	Age(P): Sex	-3.50	-84.24	77.24	0.9323	Age(P): Sex	-12.35	-96.68	71.98	0.7739	Age(P): Sex	-109.75	83.90	-1.31	0.1921
	Age(W): Sex	9.35	-82.57	101.26	0.8418	Age(W): Sex	-139.94	-243.98	-35.91	0.0084	Age(W): Sex	-571.04	-1045.93	-96.15	0.0186

Dur

		Noisy Est.	CI 2.5%	CI 97.5%	P		Periodic Est.	CI 2.5%	CI 97.5%	P		Mixed Est.	CI 2.5%	CI 97.5%	P
Dur	Intercept	1.24	1.13	1.35	<2e-16	Intercept	1.58	1.51	1.65	<2e-16	Intercept	2.02	1.82	2.21	<2e-16
	Age (P)	-0.73	-0.92	-0.54	2.28E-13	Age (P)	-0.68	-0.77	-0.59	<2e-16	Age (P)	-0.99	-1.33	-0.65	2.73E-08
	Age (W)	-0.53	-0.77	-0.28	2.75E-05	Age (W)	-0.18	-0.44	0.08	0.1789	Age (W)	-1.09	-3.22	1.04	0.315
	Sex	0.41	0.16	0.66	0.0011	Sex	0.73	0.38	1.08	4.26E-05	Sex	0.29	-0.44	1.02	0.439
	Age(P): Sex	-0.55	-0.86	-0.24	0.0006	Age(P): Sex	-0.82	-1.18	-0.46	8.13E-06	Age(P): Sex	-0.29	-1.14	0.56	0.504
	Age(W): Sex	-0.46	-0.81	-0.10	0.0119	Age(W): Sex	-0.86	-1.30	-0.42	0.0001	Age(W): Sex	-0.33	-2.76	2.11	0.792
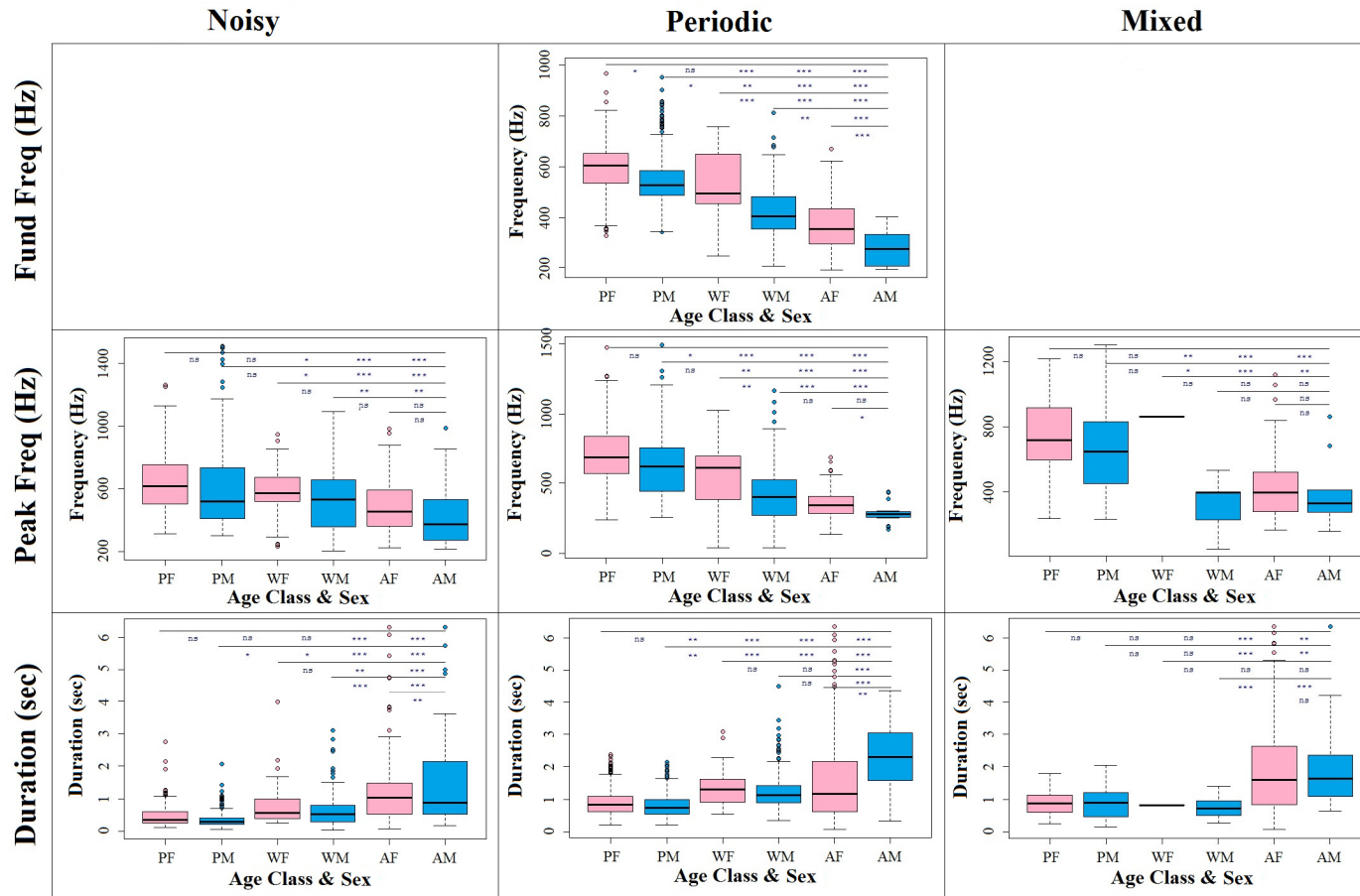
Figure 3.4. Tukey's boxplots for average fundamental frequency, peak frequency and duration of calls for age and sex classes by call type. Age was categorized into three categories; nursing pups 'P' (stages 1-4), weaned pups 'W' (stage 5 to 1 year) and adults (older than 1 year). Model contrast are shown above the graphs with ***p<0.0001, **p<0.001, *p<0.01 and not significant 'ns'. Mean and standard deviation for each parameter by call type, age class and sex is shown in Appendix III.

3.4.2 Phylogenetic Principal Components Analysis

Other call parameters also appear to vary by age and sex. Phylogenetic PCA was used to reduce the number of acoustic parameters into a smaller number of components. The Kaiser-Meyer-Olkin (KMO) measure confirmed that sample size was adequate for PCA analysis for all three calls types (for 'noisy' = 0.603, 'periodic' = 0.861 and 'mixed' = 0.817). Bartlett's test of sphericity indicated that correlations between variables were sufficiently large for all three call types ($p<0.0001$). Out of the 84 ('noisy' and 'mixed' calls) and 93 ('periodic' calls) measured acoustic parameters, ten components were kept for each call type. These ten components explained 63.01% of the variance for 'noisy' calls, 55.81% for 'periodic' calls, and 58.55% for 'mixed' calls. While ten components were retained, the majority of variance was explained by the first three components (44.9%, 39.08% and 39.96%, respectively).

A full list showing factor loadings for all acoustic variables by call type is in Appendix IV ('noisy'), V ('periodic') and VI ('mixed' calls). Although the first ten components were included to provide more information on how the call parameters vary, only the loadings for the first three components are discussed in depth here as they contributed to the majority of the data's variance. Factor loadings for the acoustic variables that contributed the most (90% of the weight) to each component are discussed (Table 3.7).

For 'noisy' calls, component 1 was negatively related to entropy and peak frequency bandwidth, and positively related to the harmonic to noise ratio. Component 2 was negatively related to the mean number of frequency peaks and entropy in the second quartile and positively related to several frequency parameters. Component 3 was negatively related to several frequency parameters at the start of the call, and positively related to several frequency parameters and the harmonic to noise ratio at the start of the call.

For 'periodic' calls, component 1 was negatively related to mean entropy, and positively related to the median minimum frequency and mean first frequency peak. Component 2 was negatively related to median and mean entropy, mean number of frequency peaks, and median maximum frequency. Component 2 was positively related with the mean harmonic to noise ratio.

Table 3.7. Phylogenetic PCA loadings for parameters that explained the majority (90%) of weight on the first three components by call type. Complete loadings for all parameters are shown in Appendix IV ('noisy'), V ('periodic') and VI ('mixed' calls).

Noisy	
PC1: Negative	
Entropy Mean	-0.899
Entropy Med	-0.898
Peak BW Medium	-0.851
Entropy Max	-0.836
Peak BW Mean	-0.832
Entropy Centre	-0.82
Entropy ¾	-0.805
PC1: Positive	
HNR Max	0.434
HNR Mean	0.522

Noisy	
PC2: Negative	
# Peaks Mean	-0.431
Entropy Centre	-0.364
PC2: Positive	
Min Freq Mean	0.602
Max Freq Mean	0.607
Peak Freq Mean	0.628
Peak Freq Centre	0.649
Min Freq Centre	0.661
Peak Freq Med	0.676

Noisy	
PC3: Negative	
Peak Freq Start	-0.722
Max Freq Start	-0.718
Min Freq Start	-0.699
PC3: Positive	
Min Freq Centre	0.22
Peak Freq Centre	0.239
Min Freq Max	0.243
Peak Freq Max	0.252
Max Freq Centre	0.258
HNR Start	0.259
Max Freq Max	0.27

Periodic	
PC1: Negative	
Entropy Mean	-0.656
PC1: Positive	
Min Freq Med	0.744
1st Freq Peak Mean	0.804

Periodic	
PC2: Negative	
Entropy Med	-0.748
# Peaks Mean	-0.722
Max Freq Med	-0.7
Entropy Mean	-0.677
PC2: Positive	
HNR Mean	0.572

Periodic	
PC3: Negative	
Total BW Medium	-0.727
Total BW End	-0.692
Total BW ¾	-0.676
Total BW Start	-0.669
Total BW ¼	-0.654
PC3: Positive	
HNR SD	0.357

Mixed	
PC1: Negative	
Entropy Mean	-0.918
Entropy Med	-0.893
# Peaks Mean	-0.843
PC1: Positive	
HNR Mean	0.486
Min Freq Med	0.515
1st Freq Peak Mean	0.547

Mixed	
PC2: Negative	
Peak Freq Mean	-0.933
Max Freq Mean	-0.933
Min Freq Mean	-0.926
PC2: Positive	
HNR Centre	0.166
HNR ¼	0.169
Entropy SD	0.177
HNR Med	0.202
HNR Mean	0.227

Mixed	
PC3: Negative	
Max Freq End	-0.65
Peak Freq end	-0.647
Min Freq End	-0.6
PC3: Positive	
HNR ¾	0.277
# Peaks ¼	0.278
HNR End	0.289
HNR Mean	0.342

Component 3 was negatively related to the peak frequency bandwidth and was positively related to the standard deviation of the harmonic to noise ratio.

For 'mixed' calls, component 1 was negatively related to entropy and the mean number of peak frequencies and was positively related to the harmonic to noise ratio, minimum frequency and the first peak frequency. Component 2 was negatively related to the mean peak, maximum and minimum frequency, and positively related to the harmonic to noise ratio and standard deviation of entropy. Component 3 was negatively related to several frequency parameters measured at the end of the call, and positively related to the harmonic to noise ratio and number of peaks at the start of the call.

3.4.3   Linear Mixed Effects Models

The ten retained components for each call type were then used as predictor variables in linear mixed effects models. A full list of the covariates used in each model by call type is shown in Appendix VII. Results of the models with the covariates of age, sex and their interaction are shown by call type in Table 3.8 ('noisy'), 3.9 ('periodic') and 3.10 ('mixed' calls). Models which retained additional covariates, such as signal to noise ratio and nonlinearities, are shown in Appendix VIII ('noisy'), IX ('periodic'), and X ('mixed' calls).

Although the first ten components were modelled to provide more information on how the call parameters varied, only the models for the first three components are focused on in detail here as they contributed to the majority of the data's variance from the PCA. For 'noisy' calls, component 1 (which was related to entropy, peak frequency bandwidth and harmonic to noise ratio) decreased with age. Component 2 (which was related to number of frequency peaks, entropy, and minimum, maximum and peak frequency) did not significantly vary for any of the covariates. However, when the covariate 'nonlinearities' was modelled as a categorical variable by type of nonlinearity (as opposed to presence / absence) the component was associated with the presence of harmonics (Appendix XI). Component 3 (which was related to harmonic to noise ratio, peak, minimum, and maximum frequency at the start and second quartile of the call) decreased with signal to noise ratio, and

Table 3.8. Linear mixed effects models for 'noisy' calls showing how PCA components varied with age (in days) and sex. Individual animal was the random effect, and when retained during model selection date was nested within individual. For sex, males were referential in the model. Significant (p < 0.05) variables are shown in bold.

	Intercept				Age				Sex				Age:Sex			
		CI				CI				CI				CI		
PC	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	5.110	0.024	10.238	0.0536	-2.743	-4.760	-0.737	**0.0087**	0.490	-5.837	6.628	0.8785	-1.104	-4.318	2.221	0.5099
PC2	0.191	1.647	2.041	0.8550	-0.177	-0.844	0.492	0.6390	-0.479	-3.173	2.212	0.7550	0.04	-1.185	1.259	0.9530
PC3	1.031	-0.935	2.993	0.3397	-0.036	-0.638	0.566	0.9127	-0.611	-3.248	2.028	0.6618	0.686	-0.368	1.671	0.1986
PC4	3.476	2.559	4.378	**0.0000**	-0.254	-0.564	0.058	0.1140	-0.331	-1.525	0.846	0.5880	0.397	-0.22	1.015	0.2160
PC5	2.977	1.422	4.525	**0.0013**	-0.408	-0.925	0.11	0.1420	-0.104	-2.268	2.051	0.9280	-0.281	-1.133	0.547	0.5250
PC6	-1.133	-2.235	-0.053	0.0736	0.613	0.295	0.909	**0.0145**	0.788	-0.401	1.903	0.2817	-0.557	-1.137	0.01	0.0830
PC7	-0.22	-2.141	1.667	0.8280	0.948	0.186	1.709	**0.0186**	-0.568	-3.093	2.035	0.6756	0.35	-0.863	1.503	0.5703
PC8	0.127	-0.757	0.997	0.7950	0.177	-0.141	0.493	0.3210	0.074	-1.212	1.379	0.9170	0.069	-0.479	0.664	0.8180
PC9	-0.458	-2.344	1.421	0.6483	0.505	-0.225	1.237	0.1888	1.85	-0.681	4.487	0.1739	-0.382	-1.546	0.73	0.5169
PC10	0.569	-0.143	1.282	0.1590	-0.106	-0.358	0.145	0.4410	0.322	-0.687	1.329	0.5630	-0.318	-0.705	0.073	0.1390

Table 3.9. Linear mixed effects models for 'periodic' calls showing how PCA components varied with age (in days) and sex. Individual animal was the random effect, and when retained during model selection date was nested within individual. For sex, males were referential in the model. Significant (p < 0.05) variables are shown in bold.

| | Intercept | | | | Age | | | | Sex | | | | Age:Sex | | | |
| | | CI | | | | CI | | | | CI | | | | CI | | |
PC	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	-10.86	-14.314	-7.391	**1.11E-05**	0.798	-0.401	1.973	0.2061	1.912	-3.26	6.784	0.422	-1.698	-3.843	0.737	0.1380
PC2	-5.108	-7.601	-2.685	**0.0006**	1.461	0.53	2.423	**0.0053**	1.155	-2.439	4.892	0.5508	0.735	-1.562	2.923	0.5239
PC3	-2.903	-3.988	-1.829	**2.44E-07**	0.753	0.321	1.188	**0.0007**	5.51	3.822	7.199	**4.59E-10**	-3.037	-4.08	-1.991	**2.11E-08**
PC4	-0.372	-1.807	1.077	0.634	0.346	-0.131	0.821	0.1882	0.328	-1.523	2.187	0.742	-0.695	-1.409	0.0216	0.072
PC5	1.042	-0.086	2.176	0.0747	-0.276	-0.725	0.173	0.2341	1.572	-0.139	3.268	0.0741	-2.21	-3.288	-1.135	**7.70E-05**
PC6	2.658	1.611	3.786	**7.17E-05**	0.971	0.522	1.366	**4.53E-05**	1.166	-0.421	2.671	0.1631	-0.788	-1.696	0.183	0.1068
PC7	2.115	1.156	3.081	**0.0004**	0.106	-0.191	0.391	0.445	-0.087	-1.307	1.146	0.8694	0.099	-0.382	0.527	**6.41E-01**
PC8	-0.617	-1.056	-0.169	**0.0083**	0.019	-0.123	0.16	0.7935	-0.385	-0.949	0.181	0.1917	0.536	0.237	0.834	**0.0008**
PC9	-0.612	-1.586	0.407	**2.60E-01**	0.935	0.524	1.309	**2.50E-05**	2.392	0.934	3.828	**0.0029**	-2.221	-3.089	-1.318	**2.07E-06**
PC10	0.439	-0.507	1.354	0.4001	0.197	-0.159	0.562	0.3183	1.649	0.249	2.972	**0.0302**	0.784	-0.069	1.595	0.0716

Table 3.10. Linear mixed effects models for 'mixed' calls showing how PCA components varied with age (in days) and sex. Individual animal was the random effect, and when retained during model selection date was nested within individual. For sex, males were referential in the model. Significant (p < 0.05) variables are shown in bold.

PC	Intercept Coeff	CI 2.5%	CI 97.5%	P	Age Coeff	CI 2.5%	CI 97.5%	P	Sex Coeff	CI 2.5%	CI 97.5%	P	Age:Sex Coeff	CI 2.5%	CI 97.5%	P
PC1	-3.086	-8.818	3.048	0.3336	-0.632	-1.814	0.558	0.3541	-1.838	-7.468	3.704	0.561	0.501	-2.059	2.987	0.7186
PC2	0.756	-6,263	8	0.8384	0.291	-1.612	2.175	0.7761	-7.745	-15.446	1.246	0.0783	5.45	0.078	10.056	**0.039**
PC3	0.648	-2.951	4.081	0.7237	0.034	-0.597	0.664	0.9209	1.103	-1.767	3.977	0.4778	-0.243	-1.448	0.951	0.7054
PC4	0.685	-0.839	2.194	0.403	-0.153	-0.65	0.346	0.568	-0.664	-1.947	0.621	0.339	NA	NA	NA	NA
PC5	-0.941	-2.196	0.303	0.1656	0.27	-0.142	0.684	0.2243	1.291	0.214	2.378	**0.0299**	NA	NA	NA	NA
PC6	1.759	-0.096	3.627	0.0899	-0.202	-0.768	0.363	0.5109	-0.261	-2.832	2.325	0.8512	-0.004	-1.091	1.063	0.995
PC7	0.269	-0.672	1.199	0.596	-0.932	-0.513	0.119	0.25	0.242	-0.587	1.645	0.587	NA	NA	NA	NA
PC8	-1.142	-4.212	2.964	0.5218	1.089	-0.104	1.973	**0.0311**	1.952	-2.297	5.614	0.3542	-0.498	-2.674	1.906	0.6857
PC9	3.162	0.946	5.341	**0.0059**	-0.133	-0.341	0.085	0.2957	0.603	-0.385	1.597	0.3034	-0.543	-1.086	-0.003	0.0663
PC10	-3.573	-5.221	-1.61	**0.0007**	1.597	0.814	2.214	**0.0001**	4.562	2.097	7.709	**0.0062**	-1.948	-3.817	-0.501	**0.0443**

increased with the presence of nonlinearities. Of the additional seven components, 4, 5, 6 and 7 were primarily related to the signal to noise ratio and the other models had no significant covariates (Table 3.8, Appendix VIII).

For 'periodic' calls, component 1 (which related to average entropy, first peak frequency and the minimum frequency) increased with signal to noise ratio and the presence of nonlinearities, and decreased with an interaction between age and signal to noise ratio. Component 2 (which was related to entropy, number of frequency peaks, maximum frequency and harmonic to noise ratio) increased with age and an interaction between age and signal to noise ratio. Component 3 (which was related to total frequency bandwidth and the standard deviation for harmonic to noise ratio) increased for age, an interaction between age, sex and signal to noise ratio, and was higher in females than males.

Although the results are limited here to report the model results of the first three components, component 5 notably followed predictions of source filter theory for peak and fundamental frequency. Component 5 (which related to the standard deviation and fourth quartile minimum, maximum and peak frequency, the fundamental frequency and the first peak frequency) was negatively related to signal to noise ratio, age interacting with sex, and sex interacting with signal to noise ratio and was positively related with the presence of nonlinearities (Table 3.9, Appendix IX, Figure 3.5). Many of the other components also related to age and sex; notably duration and frequency parameters at the start of the call increased with age (component 6), and frequency parameters and energy were higher in female calls (component 7, 8, 9 and 10) (Table 3.9, Appendix IX).

For 'mixed' calls, component 1 (which related to entropy, number of frequency peaks, first frequency peak, harmonic to noise ratio and minimum frequency) increased with signal to noise ratio, and decreased with the presence of nonlinearities. Component 2 (which related to average minimum, maximum and peak frequency, harmonic to noise ratio and entropy) increased with an interaction between age and sex, age and signal to noise ratio, and sex and signal to noise ratio. Component 2 decreased with signal to noise ratio and the interaction between age, sex and signal to noise ratio. Component 3 (which
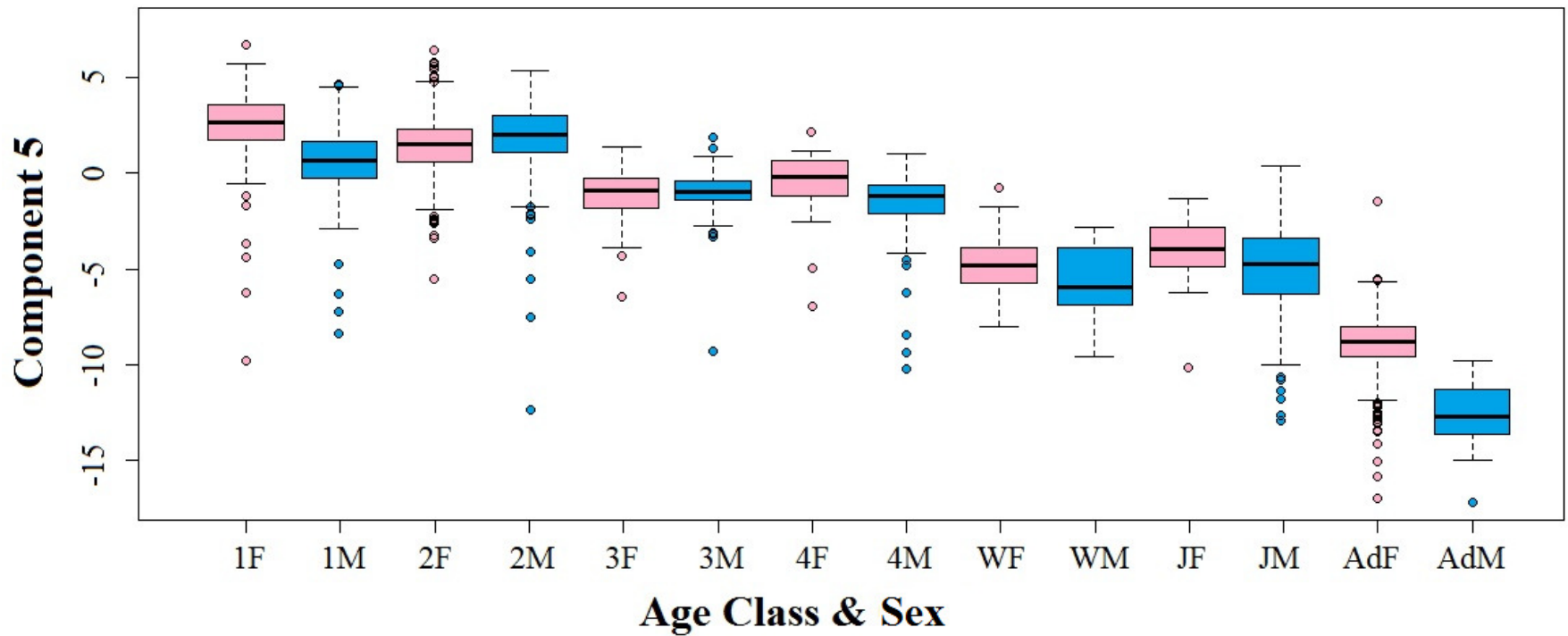
Figure 3.5. Tukey's boxplots from PCA component 5 for 'periodic' calls by age and sex. Age is indicated using Kovac and Lavigne's (1986) 5 point scale to classify pup ages. Stage 5 pups are classified as weaners 'W' (21 to 50 days old). Older animals were classified as juveniles 'J' (51 to 365 days old) and adults (older than 366 days old). Sex is indicated by 'F' for female, and 'M' for male. This component was related to the standard deviation and fourth quartile minimum, maximum and peak frequency, the fundamental frequency and the first peak frequency. These parameters decreased with age interacting with sex, signal to noise ratio and sex interacting with signal to noise ratio.

related to minimum, maximum and peak frequency, number of frequency peaks, and harmonic to noise ratio) decreased with signal to noise ratio. The other components indicated that frequency and bandwidth was higher in females than in males (component 5 and 10), and that duration and the first peak frequency decreased with age (component 8) (Table 3.10, Appendix X).

3.4.4 Cosinor and Sequence Analysis

The cosinor analysis indicated that for the most part, the seals did not exhibit diurnal variation in calling behaviour (Table 3.11). However, 'periodic' and 'mixed' underwater calls did show diurnal variation, with the majority of calls being produced between 20:00 and 22:00 each day (Figure 3.6). For the underwater 'sequence' calls, a total of 100 sequences were recorded, 54 of which were unique (i.e. never repeated). The same sequence never repeated more than nine times, and repeated sequences were usually short (up to three calls long). The sequences ranged between 2 to 24 calls in length (on average 5.146, ± 4.536), with an average intercall interval of 0.295 seconds (± 0.323). Sequences were most likely (67%) to start with the 'noisy' unit. The transition rate between calls was roughly equal going from 'noisy' to 'noisy' or to 'periodic' calls (53.1% and 46.9%, respectably). However, it was more likely to transition from a 'periodic' to a 'noisy' call (72.9%) than from 'periodic' to another 'periodic' call (27.1%).

Table 3.11. Results of the cosinor models examining if call rate follows a diurnal pattern. In-air and underwater calls were examined by total number of calls and call type. Underwater calls had an additional unique call type, 'sequences', when calls were produced with an intercall interval of less than 2 seconds. Significant ($p < 0.05$) variables are shown in bold. In models that fit sine wave predictions, both amplitude (Amp) and acrophase (Acr) must be significant.

In-air Calls			CI		
		Estimate	2.50%	97.50%	P
Total	**Intercept**	13.144	5.969	20	**0.0003**
	Amp	17.601	7.328	27.873	**0.0008**
	Acr	-0.46	-1.004	0.084	0.0971
Noisy	**Intercept**	5.739	2.954	8.524	**0.0001**
	Amp	6.916	2.923	10.908	**0.0007**
	Acr	-0.452	-0.989	0.084	0.0984
Periodic	**Intercept**	7.035	1.818	12.251	**0.0082**
	Amp	10.121	2.644	17.598	**0.008**
	Acr	-0.453	-1.139	0.234	0.1962
Mixed	**Intercept**	0.212	-0.041	0.465	0.1004
	Amp	0.444	0.077	0.81	0.0177
	Acr	-0.383	-1.134	0.367	0.3165

Underwater Calls			CI		
		Estimate	2.50%	97.50%	P
Total	**Intercept**	10.572	6.696	14	**0**
	Amp	6.299	1.22	11.476	**0.0171**
	Acr	-0.843	-1.724	0.037	0.0604
Noisy	**Intercept**	2.534	1.004	4.063	**0.0012**
	Amp	1.662	-0.355	3.68	0.1064
	Acr	-0.911	-2.242	0.419	0.1794
Periodic	**Intercept**	4.331	2.519	6.143	**0**
	Amp	4.194	1.732	0.657	**0.0008**
	Acr	-0.749	-1.358	-0.141	**0.0157**
Mixed	**Intercept**	0.142	-0.049	0.332	0.1452
	Amp	0.23	-0.013	0.474	**0.0633**
	Acr	1.39	0.161	2.619	**0.0267**
Sequence	**Intercept**	3.566	2.076	5.056	**0**
	Amp	0.338	-1.559	2.234	0.7271
	Acr	-1.141	-7.698	5.416	0.7331
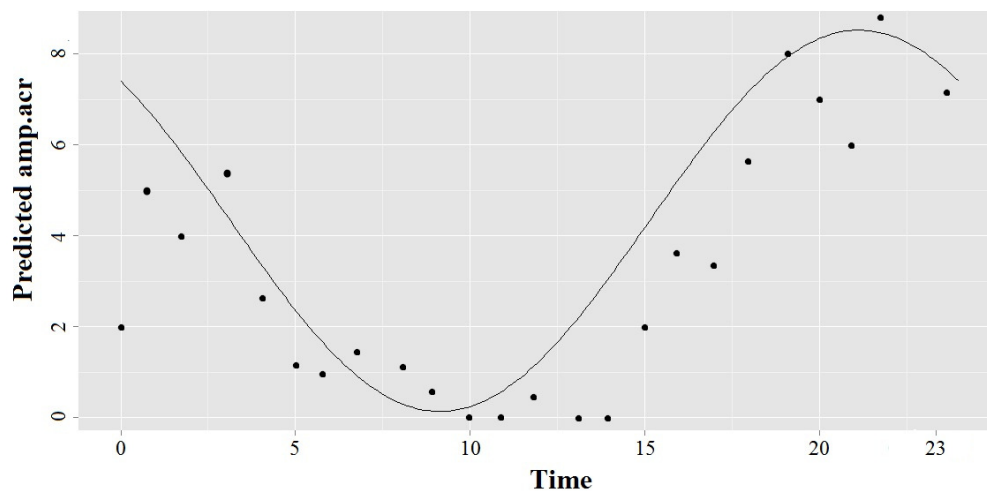
Figure 3.6. Observed call rates, averaged per hour (points) and predicted cosinor values (solid line) from underwater 'periodic' calls. Call rate followed a diurnal, sine wave pattern, with the peak number of calls being produced between 20:00 and 22:00.

3.5 Discussion

Here I have shown that grey seal calls changed with age and varied between sexes. These changes followed source filter theory predictions similarly to many other mammals; in general frequency parameters decreased and duration increased with age, and males had lower frequency and longer duration calls than females. Other acoustic parameters also appeared to vary by age and sex. The total bandwidths of the calls increased with age and were larger in females than in males. Parameters that indicated how 'noisy' a call was (i.e. entropy, harmonic to noise ratio and bandwidth of the peak frequency) decreased with age. Similarly, the variability (as indicated by standard deviation across the call) of the peak frequency and harmonic to noise ratio decreased with age.

These changes were apparent for some call types and stages of development but not others. When animals were categorically grouped by age (nursing pups, weaned pups or adults), parameter differences were apparent in some age and sex groups, but not others, depending on call type. For example, the fundamental frequency of 'periodic' calls differed between all age classes and sexes, with the exception of nursing and weaned female pups. Similarly, there was no difference in duration of 'noisy' calls in nursing and weaned pups,

106

but adults produced longer calls. The peak frequency of 'mixed' calls also did not differ between sexes, and there was only a difference between nursing pups and adults.

Some of these differences may be attributed to methodological limitations. For example, the sample size for some age classes of 'mixed' calls was very small and may not be representative. However, other differences may be biologically relevant. Grey seal mothers nurse their pups for approximately three weeks before weaning, at which time they leave their pup on land and return to the sea (Kovacs & Lavigne, 1986). There are several potential cues which could influence when the mother chooses to wean her pup, and if pup call parameters vary between nursing and weaning then acoustic cues could be used. Similar acoustic cues could be used to gauge maternal investment in pups; some evidence suggests that grey seals differentially invest in male pups (Kovacs & Lavigne, 1986; Anderson & Fedak, 1987). However, other studies indicate there is no difference in maternal investment between sexes (Bowen et al., 1992; Smiseth & Lorentsen, 1995). Although acoustic cues may be reliable indicators of age class and sex, it is possible that the seals do not use the signals. Future research could test this by utilizing playback techniques to further investigate the role of pup call characteristics on maternal care.

While past research has shown varying trends in pinniped vocal development, my results indicate that grey seal calls decrease in frequency and increase in duration and vary between sexes similarly to most mammals. The variable trends observed in previous pinniped studies may be explained by the differing growth rates between species, or methodological differences such as the developmental sampling period. As pinniped species show variable growth and development, it is difficult to compare between species, particularly when different periods are examined. For example, some studies have focused on changes in calls during the nursing period (Job et al., 1995; van Opzeeland & van Parijs, 2004; Khan et al., 2006; Muñoz et al., 2011), some extended to weaning period (Charrier et al., 2003), while others examined juveniles into adulthood (Davies et al., 2006; Rogers, 2007; Sanvito et al., 2008). This presents difficulty when trying to examine developmental trends across species

with varying growth trends, especially with so few studies. Further studies examining call development, and corresponding physical growth, in pinnipeds would be valuable.

Alternatively, the variable trends observed may be attributed to how call parameters were measured. Although in general my results show that frequency parameters decreased with age, there were exceptions. For example, as shown by component 6 of 'periodic' calls, frequency at the start of the call increased with age contradictory to prediction from source-filter theory and the trends observed in the majority of the data. This difference could potentially be attributed to other factors which may impact the initial generation of a sound, such as motivational state or physical constraints. If developmental patterns differ depending on where parameters are measured within the calls, perhaps the variable trends between studies can be attributed to methodological differences. Future research would benefit from standardized parameter measurements to facilitate comparison between studies.

Ideally, studies examining call development would record all animals in the same, controlled environment. This study was limited in that recordings were made in varying outdoor and, occasionally, very noisy environments. To help control for variability in recording quality, the signal to noise ratio was used as a covariate in the models. This appeared to be relatively successful; when fitting the optimal models, signal to noise ratio was retained as a covariate in 23 out of 30 models. Some of the trends observed in the data would not have been apparent if the interaction between age or sex and signal to noise ratio was not taken into account. I do not believe that the observed age and sex trends can be attributed to recording quality, but it is possible that recording in a noisy environment masked additional changes that may occur and were not detected. If possible, future investigations of call development should attempt to use more controlled recording conditions.

Here, I modelled changes in very specific acoustic parameters (fundamental and peak frequency, and duration) based on source filter theory predictions. Additionally, I looked at several different parameters measured at several locations across the calls by using phylogenetic PCA to narrow

parameters into a small number of components. My analysis showed that frequency and energy parameters explained the majority of the variability in the calls, particularly when measured at the start, middle, minimum, maximum, mean, and standard deviation of the call. However, using so many variables may have introduced noise into the analysis, and masked some developmental patterns. While our results are thus limited, future research can benefit from our findings to target specific acoustic variables and measurement locations of interest.

While previously grey seal pups have been described as having only one call type, the periodic, harmonic 'mother attraction' call (Asselin et al., 1993; Caudron et al., 1998; McCulloch et al., 1999), other studies investigating vocal development in other species examine multiple, distinctive call types (e.g. wolves, Coscia et al. 1991; marmosets, Pistorio et al., 2006; elephants, Stoeger-Horwath et al. 2007). Here, I primarily classified calls into 'periodic', 'noisy' and 'mixed' call types. All three of these call categories were present in the grey seal pups within 48 hours of birth. As call categorization is difficult to compare across grey seal studies due to inconsistencies in human classification, this study was unable to examine when previously described call types in adult seal repertoires appeared within the pups. However, it is clear that past examinations of the grey seal pup repertoires have been too simplistic; grey seal pups produce more varied calls than the previously described 'mother attraction' call (Asselin et al., 1993; Caudron et al., 1998; McCulloch, 1999; McCulloch et al., 1999). Future investigations would benefit from examination of pup repertoire development considering multiple call types. I suggest this should be done by classifying calls using broader types, as I have done here, as opposed to the small, inconsistently applied categorization previously used. A more standardized call type classification of grey seal repertoires, examining how call type usage changes throughout development, would be beneficial.

It is also unknown as to how grey seals may perceive these call types; although they are capable of discriminating between these types (see chapter 4), they may naturally perceive additional categories. Future research should consider using cluster analysis techniques to investigate how call classes may

be discriminated by acoustic parameters. This should be done in conjunction to playback studies to determine if the animals perceive these differences between call groups, and behavioural observations investigating if such call classes are relevant to the animals' behavior.

In addition to these three main call types, underwater 'sequence' calls were also examined. The captive juvenile seals produced sequences of 'noisy' and 'periodic' calls that were consistent with previous descriptions of 'rup' and 'rupes' (Asselin et al., 1993; McCulloch et al., 1999). The order of the calls was relatively flexible, although a 'periodic' unit was more likely to be followed by a 'noisy' unit. The sequences were also variable in length and intercall duration. There is some question as to the categorization and use of these calls; previous descriptions of the 'rup' and 'rupe' make it appear as though the 'rup' may be the 'noisy' unit of the 'rupe' (Asselin et al., 1993; McCulloch, 1999). Given this variability, further description of the underwater call repertoire and usage, particularly for multiple element calls, would be valuable. Future research should consider modelling call structure, providing more description as to how to classify sounds into subunits, calls and series in conjunction to examination of the animal's use and classification of vocalisations.

The current study was limited in that the captive juveniles did not vocalise as often as the pups. This could be due to age related changes; in other species call rate decreases with age, particularly after weaning, presumably because the young no longer need to elicit maternal care (Motomura et al., 2002; Mandelli & Sales, 2004). Alternatively, the seals may have vocalised less frequently in captivity; bearded seal pups raised in captivity never vocalised until the males reached sexual maturity, and while the captive vocalisations resembled wild calls, they showed fewer call types (Davies et al., 2006). As the seals vocalised less frequently in captivity (where all of the underwater recordings were made) fewer calls were recorded and it was not possible to analyse the development of underwater calls separately from in-air calls. Future studies could obtain a larger number of samples by recording wild populations to further investigate the underwater repertoire and call usage. Similarly, the pups that had been followed from birth produced too few sounds as juveniles in

captivity to examine changes in an individual's repertoire over time or with physical growth. Future studies may benefit from using acoustic tags, paired with periodic growth measurements, to record call development of individual weaned pups in the wild.

This study had several limitations that should be kept in mind when interpreting the results. Although the adult females could be reliably identified using unique coat patterns, the pup's identity was assumed based on their association to their mother, approximate age and location. While this assumption is most likely true, some cases of allo-suckling have been observed in grey seals (McCulloch et al., 1999) and it is possible that pups could have been misidentified. Similarly, males could not be individually identified with coat patterns. Thus, males could only be recorded from single observation sessions and were not identified across days. Additionally, recorded vocalisations were attributed to individual animals using video footage based on corresponding visible body movements. Certain call types may have been more likely to be made with movements undetectable by the video footage. In cetaceans vocalizing animals are sometimes identified using body movements and bubble streams (e.g. McCowan & Reiss, 1995). However, this method has since been found to be unreliable as it is biased towards certain call types (Fripp, 2005). While this is less of a concern as the in-air sounds produced by seals appear to correspond to relatively large movements, potentially quiet sounds produced with the mouth closed, such as 'humming' sounds, may have been excluded from our analysis. Thus, our method may have resulted in biased sampling of the seal's call repertoire.

Human interaction may have also resulted in unrepresentative repertoire sampling. On the Isle of May, most of the recordings were made with the researcher out of sight or at least 10 m away from the animals. This human disturbance was deemed as being minimal because the animals appeared to display normal behaviours such as nursing, sleeping, etc. when humans were present. However the animals were still likely to be aware of human presence, which may have altered natural calling behaviours.

This may have similarly occurred in captivity; human presence may have made the seals more or less likely to vocalise, influenced the call types or call structure produced. This effect may not have been constant over time; with experience the seals may have had a decreased stress through habituation or learned to associate humans with presence of food, which may have changed vocal behaviour. The cosinor test indicated that for the most part the seals' calls did not follow a diurnal pattern. Vocalisations occurring during the day may have thus been representative of their normal behavior, and may indicate that human presence was not as likely to affect the animal's vocalisations. However, for 'mixed' and 'periodic' underwater calls, the seals did show a diurnal pattern, producing the most calls between 20:00 and 22:00. This is interesting in comparison to Asselin et al. (1993) findings that only 'clicks', which would have been classified as 'noisy' calls in this investigation, showed diurnal variation. The increased production of 'clicks' at night led to the hypothesis that seals may use the calls for navigation (i.e. active echolocation), which has since been deemed false (e.g. Oliver, 1978; Schusterman, 2008). The function of these call types, and why certain types are produced predominantly at night, remains unclear. Future studies should consider such possible diurnal patterns and the influence of human interaction when recording animal repertoires.

The effect of human presence may have also been related to training experience; two of the captive seals, Clark and Zola, were trained using operant conditioning to vary number, duration and frequencies of calls to imitate novel stimuli (see chapters 4 and 5). Although calls recorded during training sessions were not used in this analysis, training may have affected the calls produced outside of sessions as well. The use of conditioning to shape the seals' calls may have influenced call parameters; for example, they may have been more likely to produce calls with frequencies that had been most heavily rewarded. Training may also have changed call rate; to ensure the seals had learned the task vocalisations were cued and put under stimulus control. This entails that the seals only produce the sounds on cue, and are silent when the cue is not present. The lack of this cue outside of training sessions could have resulted in the decreased calling rate observed in the captive juvenile seals. This does not

seem likely however; the juvenile animals who were trained produced more calls than those who were not. Thus, training may have alternatively made the juvenile seals more likely to call, and perhaps natural calling rates of juveniles would be lower.

An additional experiment may have altered the seals calling behaviour. Three of the weaned seals took part in a study investigating the effects of oxytocin during social interactions (Robinson, 2014). Oxytocin is associated with social bonding (such as mother/pup contact), and pups who received oxytocin stayed in closer proximity and exhibited fewer aggressive behaviours to each other seals in comparison to pups who received saline (Robinson, 2014). As relatively few calls were recorded during these trials, they were not separately analysed in this study. Although only a small sample size was examined, the developmental changes of these animals appeared to be consistent despite oxytocin manipulations. However, it is anticipated that pups who received oxytocin would have altered vocal behavior. Oxytocin might be expected to alter calling behaviour by favouring certain call types, such as those associated with more affiliative social interactions. Future research could further explore the effects of oxytocin on call behavior by examining a larger sample of animals.

Our results have shown that grey seal pup call development is more complex than has been previously described and follows trends similar to many other mammals. However, it is still unknown as to how sound production mechanisms change with development. Future research considering morphological growth of the vocal tract and its correlation to call changes would be valuable. Similarly, it is still unknown as to how an individuals' repertoire develops, and to what extent learning impacts call usage, production and comprehension. Further studies on learning would be additionally valuable when examining how individual animals develop and use call repertoires.

Chapter 4
Vocal Usage Learning in the Grey Seal:
Generalization of Call Signal Classes

4.1   Summary

Past research found that grey seals (*Halichoerus grypus*) were capable of vocal usage learning as they were able to classify sound signals into call types using vocal responses (Shapiro et al., 2004). However, they were only able to do so using a trained set of stimuli, and were unable to generalize to novel exemplars. Here I trained a juvenile grey seal to discriminate novel calls into two classes, 'growls' and 'moans', by vocally matching call types (i.e. the seal moaned when played a moan and growled when played a growl). My method differed from the previous study as I trained the animal using a comparatively large set of exemplars with standardized durations, consisting of both the seal's own calls and those of two other seals. The seal successfully discriminated growls and moans for both her own (94% correct choices) and other seal's (87% correct choices) calls. The seal's accuracy significantly improved across test sessions, and was higher during the first presentation of a sound from its own repertoire, but decreased after multiple exposures. This pattern was not found for calls from unknown seals. Factor analysis for mixed data (FAMD) identified acoustic parameters that could be used to discriminate between call types and caller identity. Growls and moans differed in noise, duration and frequency, whereas individual callers only differed in frequency. This suggests that the seal could have gained information about both call type and caller identity from conspecific vocalisations.

4.2   Introduction

Vocal usage learning is the process whereby individuals learn to produce a pre-existing signal from their vocal repertoire in a new context through experience (Janik & Slater, 2000). For example, several studies have trained animals to elicit pre-existing sounds in novel testing procedures, such as a California sea lion (*Zalophus californianus*) trained to produce 'click' vocalisations to indicate when it perceived a sound to measure the animals'

behavioural audiogram (Schusterman et al., 1972). This may similarly occur in wild populations; for example food calls are used by several species to provide information to others such as the quality, quantity, and location of specific foods (e.g. Hauser, 1998; Clay et al., 2012; King & Janik, 2015). Several mammalian species produce such context-dependent calls (e.g. fallow deer, *Dama dama:* Charlton & Reby, 2010; domestic dogs, *Canis familiaris:* Taylor et al., 2009; goats, *Capra hircus:* Briefer et al., 2015). While it is likely that learning has occurred in these cases, it cannot be ruled out that call usage is innately determined to some extent.

In a previous study, Shapiro et al. (2004) examined how to test vocal usage learning abilities systematically. They identified four levels which demonstrate increasing complexity and control over vocal usage learning. The simplest skill involves the animal producing a call reliably when cued in an arbitary context, such as with a trained hand signal. Next, the animal must also remain silent in the absence of that cue such that vocalisations are only produced when given the appropriate stimulus (i.e. the behaviour was under stimulus control). Then, multiple cues are used to produce distinct call types. To accomplish this, the animal must be able to discriminate between a set of cues, and classify the calls they produce into appropriate types. Lastly, novel stimuli from trained signal classes are used to cue the individual to produce the same distinct vocalisations. This requires the individual to classify and generalize novel cues in addition to their responses (Shapiro et al., 2004).

A fundamental ability in animal cognition is the capacity for classification and generalization of stimuli. Classifying objects reduces large numbers of stimuli into a small group of categories, allowing animals to cope with stimulus variability. Generalization places new stimuli into pre-existing categories, speeding up signal processing and response time. Classification and generalization of stimuli has been well investigated in the visual domain for several species, such as bottlenose dolphins (*Tursiops truncatus:* Mercado et al., 2000), domestic dogs (*Canis familiaris:* Range et al., 2008), pigeons (Wasserman et al., 2006), parrots (e.g. Pepperberg, 1996 & 1999) and nonhuman primates (Schrier & Brady, 1987; Neiworth & Wright, 1994).

Many species are similarly capable of auditory categorization. Past research has mainly concentrated on the classification of bird song (e.g. Beecher et al., 1994) and predator-specific alarm calls (e.g. Seyfarth et al., 1980a; Greene & Meagher, 1998; Cäsar et al., 2013). Similarly, some species categorize conspecific call types by responding to another individual's calls with a matching call type in specific contexts, referred to as antiphonal calling (e.g. Ghazanfar et al., 2001; Soltis et al., 2005). Animals can also use sound for classification of objects, particularly in echolocating species (e.g. Roitblat et al., 1990; Helverson, 2004). Perhaps most importantly, auditory categorization helps animals to deal with individual variation in communication calls, where classes of signals have different meaning. Previous studies have used playbacks and discrimination training to find out how individuals categorize sounds, providing insight into how animals judge conspecific stimuli (e.g. Weary & Krebs, 1992; Lind et al., 1996; Maros et al., 2008; Candiotti et al., 2013) and/or heterospecific sounds, including human speech and music (e.g. Kuhl, 1981; Porter & Neuringer, 1984; Pepperberg, 1999; Bloomfield et al., 2003; Brooks & Cook, 2010; Hoeschele et al., 2012).

Given the importance of auditory categorization in establishing a common communication code, it would be surprising if animals were unable to generalise between calls that are acoustically similar. It would be expected that communication systems should evolve such that the production of specific signals is linked to the ability to perceive the differences between signals and interpret their meaning. It is thus surprising that grey seals (*Halichoerus grypus*) have been shown to be incapable of call type categorization of novel signals (Shapiro et al., 2004). Shapiro et al. (2004) found that two juvenile grey seals, "Kylie" and "Oscar", were capable of discriminating between two call types: 'growls' and 'moans'. These call types are easily distinguishable by human observers: growls are noisy calls with a upper bandwidth limit extending to frequencies up to 20 kHz, whereas moans are periodic calls with a harmonic structure and a upper bandwidth limit rarely exceeding 5 kHz (Shapiro et al., 2004; Figure 4.1). Kylie and Oscar were trained to vocalise upon the presentation of a sound: recordings of the seals' own growls and moans. The
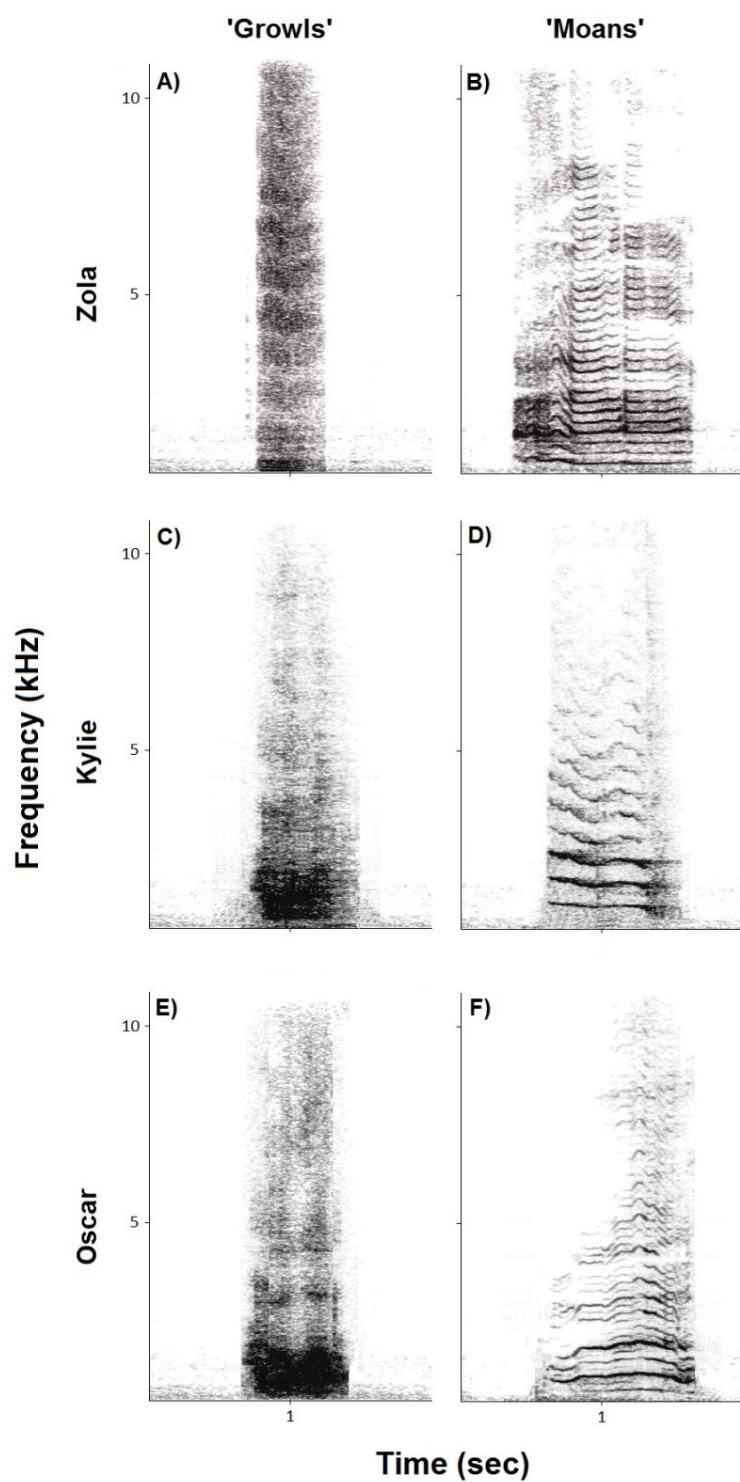
Figure 4.1. Spectrographic examples of call types from the three seals. Growls are noisy calls, and moans are periodic calls with a harmonic structure. Example growls are from A) Zola, C) Kylie and E) Oscar and moans from B) Zola, D) Kylie and F) Oscar. Spectrograms created in Avisoft-SASlab Pro (FFT size: 2048, frequency resolution: 46.87 Hz, time resolution: 10.66 ms, weighting function: hamming, window width: 100%).

seals were trained with a set of nine exemplars (five growls and four moans for Kylie, vice-versa for Oscar), and were reinforced for responding with a matching call type (i.e. responded with a moan when played a moan and a growl for a growl). Both seals successfully learned the task after several trials (~800 trials for Oscar and ~1,650 for Kylie). However, neither seal accurately responded when presented with novel exemplars (new sets of their own growls and moans). One seal responded variably to the new calls, while the other consistently responded with a growl. While the seals were capable of discriminating using a trained set of stimuli, when presented with novel sounds performance diminished. Shapiro et al. (2004) concluded that although their grey seals were able to discriminate between call types, they were unable to generalize classification to novel exemplars in this experiment.

Although this experiment failed to demonstrate the animals' abilities to classify novel stimuli, this does not necessarily indicate the seals lacked the capacity. The test paradigm in which the seals were tested, where the animals were required to indicate categorization using a vocal match, may not have been appropriate. Previous investigations documenting call categorization typically use habituation/dishabituation paradigms, which utilize natural behavioural responses to evaluate call categories (e.g. Fisher, 1998). Here, the seals were presented with natural call types, and the behavioural significance of the calls may have precluded responding with a matching call type.

There are also several methodological explanations for why the seals were unable to generalize novel exemplars, such as the training paradigm. Shapiro et al. (2004) used a small, fixed set of stimuli (nine exemplars) to train classification before testing generalization with new stimuli. Although such small training sets have been successfully used for discrimination tasks in other species (e.g. Watanabe, 2001; Brooks & Cook, 2010), it is possible that the seals simply learned through trial and error to respond with the appropriate call type for each separate stimulus as opposed to generalizing the calls into categories. This is consistent with results from other auditory classification studies which have been similarly unsuccessful at training call type categorization in rhesus monkeys (*Macaca mulatta*: Le Prell et al., 2002) and a

beluga whale (*Delphinapterus leucas*: Vergara, 2011). Both studies utilized small sets of training stimuli (for the rhesus monkeys, eight exemplars of each call type and for the beluga, three exemplars of each) before testing generalization to novel sounds. In a visual discrimination task, rhesus monkeys learned object classes faster and with greater accuracy when trained with larger exemplar sets (Schrier & Brady, 1987).

Thus, the inability to generalize may be attributed to the use of small training sets; other classification tasks have utilized large, variable sets of training stimuli. In a particularly notable example, an African grey parrot (*Psittacus erithacus*), "Alex", was successfully trained to classify objects in several ways including colour, size and shape, using dozens of training and hundreds of testing exemplars (Pepperberg, 1999). For example, during a size discrimination task Alex was required to identify which object in a pair was either larger or smaller (Pepperberg & Brezinsky, 1991). Initially, the training set consisted of eighteen exemplars but was then expanded to include thirty-two different types of objects, with exemplars of each type varying in size and colour. By using a large exemplar set Alex was unable to simply learn specific responses to specific stimuli, and had to generalize across exemplars to accomplish the task (Pepperberg & Brezinsky, 1991). Notably, during these tasks Alex was verbally cued using human speech to classify the objects, additionally demonstrating his ability to discriminate between auditory speech sounds and generalize words produced by different speakers (Pepperberg, 1999). Other species are similarly able to discriminate between hundreds of speech sounds (e.g. primates: Savage-Rumbaugh et al., 1993; dogs: Kaminski et al., 2004; sea lions: Schusterman et al., 2002; dolphins: Herman et al., 1993; parrots: Pepperberg, 2002).

Another possible explanation for Shapiro et al.'s (2004) result is that the average duration of signals changed between the training set and novel stimuli. For example, during training Oscar was played moans with an average duration of 1.5 seconds (±0.9) and growls of 2.5 seconds (±1.2). For the novel exemplars, duration of both call types decreased (moans 0.4 seconds ±0.1, and growls 0.7 seconds ±0.1). If duration was used as a cue to discriminate

between call types, this would prevent the seals from successfully performing the task.

Additionally, Shapiro et al. (2004) only presented the seals with their own sounds. While this method does test the animal's ability to classify its own calls, the seal should also be capable of generalizing novel exemplars from conspecifics. For auditory categorization to be used for communication, animals must both be able classify the calls they produce as well as the calls they perceive from others. Thus, generalization should be tested for both the seal's own calls as well as for those of other individuals.

The present study tested a juvenile grey seal's abilities to discriminate her own sounds in addition to those produced by other, unknown juvenile grey seals. I report that the seal was capable of generalizing novel exemplars of two vocalisation classes, growls and moans, when trained using a large, continually changing set of stimuli with standardized durations.

4.3   Methods

4.3.1   Subject

A juvenile female grey seal, Zola (tag numbers 73254/5), born November 7, 2011 on the Isle of May (Firth of Forth, Scotland), was the subject of this study. After weaning, Zola was transported to our laboratory (St. Andrews, Scotland) on December 6th, 2011. She was housed with four other juveniles in three enclosures including a large rectangular pool (42 m x 6 m x 2.5 m) and two circular pools (3 m x 5 m x 2 m). The seals were fed a varied diet of several fish species (mostly consisting of herring, *Clupea harengus*, and sprat, *Clupea sprattus*). Training and testing occurred at the facility for 12 months, after which Zola was released into the wild.

4.3.2   Acoustic Recordings

Acoustic recordings were obtained using a Sennheiser MKH 416 P48 directional microphone (frequency response 40-20,000 Hz ± 1 dB) and an Edirol FA-66 external sound card (sampling rate 96 kHz, 24-bit) with a laptop computer. Weather permitting, sessions were concurrently video recorded using

a Sony HDR CX250E video camera. Stimuli were played from the laptop using the FA-66 through an external Skytec 170.170 active speaker (frequency response 32-22,000 Hz). Sounds were simultaneously played, recorded and spectrographically monitored in real time using the program Audacity 1.3 (sampling rate 96 kHz, 24-bit; Audacity Team, 2012).

4.3.3 Training and Testing Procedure

Positive reinforcement behavioural training was conducted using the seal's normal daily diet as a reward. Husbandry training began in January 2012 and focused on general behaviours including exiting the water and stationing at the poolside for testing. Vocalisation training began in March 2012. Initially Zola was reinforced for any sound she produced while in the same enclosure as the other seals. This was done because the seals vocalised more frequently while in a group, and were relatively silent when isolated. Once Zola regularly vocalised, she was separated from the others for individual sessions. At this time Zola produced the two distinct call types, growls and moans, as defined by Shapiro et al. (2004). Each call type was paired with a unique hand cue. The hand cues were then faded out to be replaced with the presentation of a sound (either a growl or moan), followed by an LED light. The light was used to indicate when the sound cue had finished, at which time the seal could respond. After the hand cues were replaced by sound, all training and testing sessions occurred with the experimenter waiting quietly out of sight.

Simultaneous to training the vocal responses, Zola was reinforced for staying quiet when the vocal cues were not presented. This was shaped until she remained quietly at station between trials for up to fifteen minutes at a time. Although between most trials this stationing period was less than a minute, in some cases the inter-trial interval was extended to avoid an overlap with environmental noise. Training sessions were variable in length depending on the seal's performance; a set amount of food was used to reinforce the seal, and sessions continued until all the fish was used (mean number of trials per session = 61, SD ± 26, mean session duration = 37 minutes, SD ± 9).

Throughout training and testing, each trial was initiated by the seal leaving the water and stationing in front of the speaker (see Figure 4.2 for diagram of the testing enclosure). Once positioned, a sound (either a growl or a moan) was played. After the sound ended, the LED light illuminated and cued the seal's response. The animal's response was evaluated using visual judgement of a real-time spectrogram display comparing the played stimuli directly to the seal's response in Audacity. If correct (the seal responded with the same call type as the one played), the seal was reinforced with fish after it stopped vocalizing. If incorrect (the seal responded with a different call type than the one played), a least reinforcing stimulus (LRS) of five to ten seconds was used before the seal could initiate the next trial. An LRS is comparable to a 'time out', a set time period during training in which the animal has no opportunity to earn reinforcement. While during a 'time out' the trainer is typically removed from the session (such as by turning their back), during a LRS the trainer makes no response. Examples of correct and incorrect responses are in Appendix XIII.

Training continued until the seal had reached the criterion of seven consecutive sessions with overall accuracy above 80%, at which point testing began. Stimuli were presented in blocks of 50 trials, with each test session consisting of either 50 or 100 trials (mean session duration = 47 minutes, SD ± 7). If the seal's response overlapped part of the played stimulus, the seal was not reinforced, but the call was still used when measuring the seal's accuracy (this occurred in 136 of the 1,182 training and 72 of the 700 test trials).

4.3.4   Sound Stimuli and Acoustic Analysis

Training stimuli were composed of Zola's own previously recorded vocalisations in sets of 10 to 20 novel exemplars per session. All calls were between 0.7 and 1.2 seconds in duration. The sounds were played randomly with one exception; if Zola responded incorrectly, the sound was replayed up to three times until she produced the correct response. If after three replays Zola
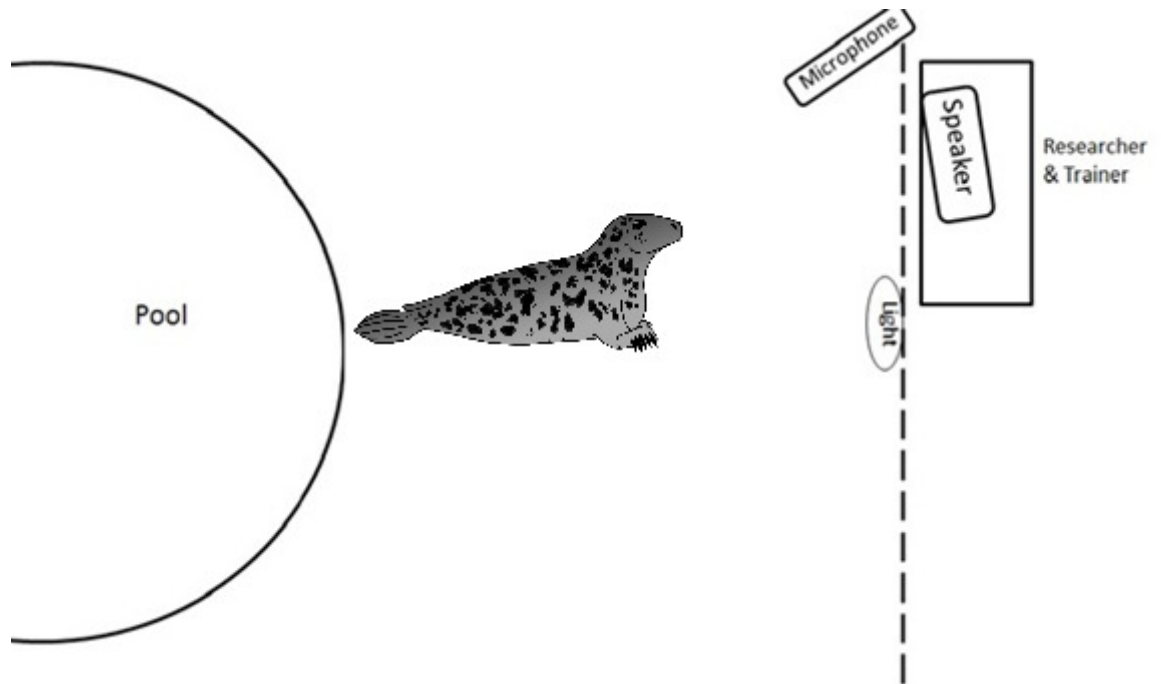
Figure 4.2. Diagram of the testing enclosure (not to scale). The seal stationed approximately one metre from the speaker and microphone. Dashed line indicates fence separating the seal from trainer, researcher and testing equipment. Grey seal image provided by Kelly Robinson.

had not responded correctly I moved on to the next call. This 'self-correction' procedure was only used during training trials, not in the test phase of the study. In total, 212 different calls (106 growls and 106 moans) were used throughout training.

Test stimuli consisted of 140 novel calls, 70 produced by Zola and 70 from the two juvenile seals studied by Shapiro et al. (2004), Kylie and Oscar, who were unknown to Zola. Again, all selected calls were between 0.7 and 1.2 seconds in duration. During each trial, one of the 140 novel call stimuli was played. Stimuli were presented in blocks of 50 trials (five moans and five growls), presented five times per stimulus in random order assigned by a Gellerman series (Gellerman, 1933). Zola was thus tested in 14 test blocks (700 trials).

Frequency and time parameters were measured using Avisoft-Saslab Pro 5.02.04 sonogram software. Parameters were chosen based on previous

studies examining grey seal vocal repertoires (McCulloch et al., 1999; Asselin et al., 1993). See Table 4.1 for a list of measured parameters and their definitions.

4.3.5   Statistical Analysis

All statistical analyses were conducted using R 3.0.1. Two independent human observers classified Zola's responses as growls or moans post hoc. At this time, the observers classified only the seal's response; they were unaware of which stimulus was played during each trial. To confirm that these calls were classified consistently, Cohen's Kappa was used to assess inter-observer reliability. To ensure human classification was not arbitrary, K-means cluster analysis (k = 2), was conducted on the frequency and time parameters measured with Avisoft-Saslab Pro (R Core Team, 2013). The human classifications were compared to the two clusters, and those that were in agreement were used to determine call classification when scoring the seal's accuracy. For the trials in which the seal's call overlapped the played signal, the call type was still identifiable and the response was included when measuring the seal's accuracy.

The seal's accuracy during training sessions was defined as the percentage of correct responses per session. A nonlinear least squares (NLS) model from the stats package (R Core Team, 2013) was used to examine how the subject's overall accuracy changed over the learning period by date and call type. As session length throughout training was dependent upon the seal's performance, the number of trials per session varied and the model was weighted by number of trials per day. The seal's overall accuracy was fitted with a polynomial regression. Terms up to the 7th order were selected using the sample size-corrected Akaike Information Criterion (AICc) score.

For test sessions, the seal's accuracy was defined as the percentage of correct responses by blocks of 50 trials, and subsequently compared between responses to its own calls and those of the two unknown seals. To determine if other parameters might have affected the seal's test performance, a generalized linear model (GLM) using the 'stats' package (R Core Team, 2013) was fitted. Accuracy was predicted using a binomial distribution and logit link

Table 4.1. Definitions for measured acoustic parameters. All parameters were measured using Avisoft (FFT size: 2048 frequency resolution: 46.87 Hz, time resolution: 10.66 ms).

Parameter	Definition
Duration	Time from the beginning to end of the call, within 35 dB of the maximum frequency spectrum.
Peak Frequency	The frequency with the highest amplitude measured at the beginning, middle, end, and overall for the whole call.
Fundamental Frequency	Measured only for periodic, harmonic calls. Measured at the centre of the call as the lowest integer multiple of corresponding amplitude peaks in the frequency spectrum.
Maximum Frequency	The highest frequency reached within 20 dB from maximum of the frequency spectrum.
Wiener Entropy	The ratio of the geometric mean to the arithmetic mean of the spectrum. This parameter is closer to zero for pure-tone signals and closer to one for random noise.
Harmonic to Noise ratio (HNR)	The dB ratio between the harmonic and non-harmonic energy. The ratio is averaged for each frequency bin (46.87 Hz), for a 10.66 ms section measured at the centre of the call.
Spectral Richness (SR)	A measure of noisy noise and harmonic spacing as defined by Miller & Murray (1995). Calls are rated on a scale of 0 to 3 based on separation between harmonic bands (H) and the width of the fundamental frequency band (F). F is measured by subtracting the lowest from the highest frequency of the fundamental at the band's widest point. H is measured by subtracting the highest frequency of the fundamental from the lowest frequency of the second harmonic at the centre of the call. A score of 0 is applied to noisy signals, 1 if $H > F$, 2 if $H = F$, and 3 $H < F$.

function with the following covariates: date, trial number, signal exposure (the number of times the seal had been exposed to stimuli; each test signal was played five times), call source (Zola, Kylie's and Oscar's calls), and call type (growl versus moan). All possible combinations were tested, including interactions between date, call source, and call type. Model averaging was performed on the 'best' models (all models within 10 AICc points of the best model, shown in Appendix XIV) using modelavg (package MuMIn, version 1.9.13: Barton, 2013) to obtain the final model.

Factor analysis for mixed data (FAMD) was used to identify cues the seal could have used to distinguish between call types (package FactoMineR, version 1.25; Husson et al., 2014). FAMD was chosen because it allowed both

continuous and categorical parameters to be used to group calls similarly to K-means cluster analysis (Ding & He, 2004). Calls were analysed both by call type (growl and moan) and by individual (Zola, Kylie, and Oscar). This determined what parameters explained variation between call types across individuals, and those that explained individual differences.

4.4   Results

4.4.1   Call Type Classification

Human classification of call types had high inter-observer agreement (99.7%, Cohen's Kappa = 0.955, SE ± 0.011). The accuracy of this classification was confirmed by the K-means cluster analysis. All calls grouped by K-means were in agreement (100%) with one of the two human's categorization, which was also in complete agreement (100%) with the experimenter who reinforced the seal during testing. Thus, this grouping was deemed objective and was used to score the seal's accuracy for the rest of the statistical analysis.

4.4.2   Training

Accuracy throughout the training period is shown by session in Figure 4.3. A polynomial trendline (order 3) shows overall performance, as determined by a best fit NLS model weighted by the number of trials per day (mean number of trials = 61, SD ± 26). Interestingly, towards the end of the training period Zola's accuracy for moans decreased. However her accuracy for growls increased, which kept her overall accuracy above criteria. Zola reached criterion (seven consecutive sessions > 80% accuracy) after 1,182 trials across 18 sessions. This fell within the time needed by Kylie (~1,650 trials, 47 sessions) and Oscar (~800 trials, 23 sessions) to learn the task (Shapiro et al., 2004).

4.4.3   Testing

Accuracy throughout testing is shown by blocks of 50 trials in Figure 4.4. Overall, Zola responded with 93.7% accuracy to her own calls and with 87.2% accuracy to Kylie and Oscar's calls. This performance was significantly better
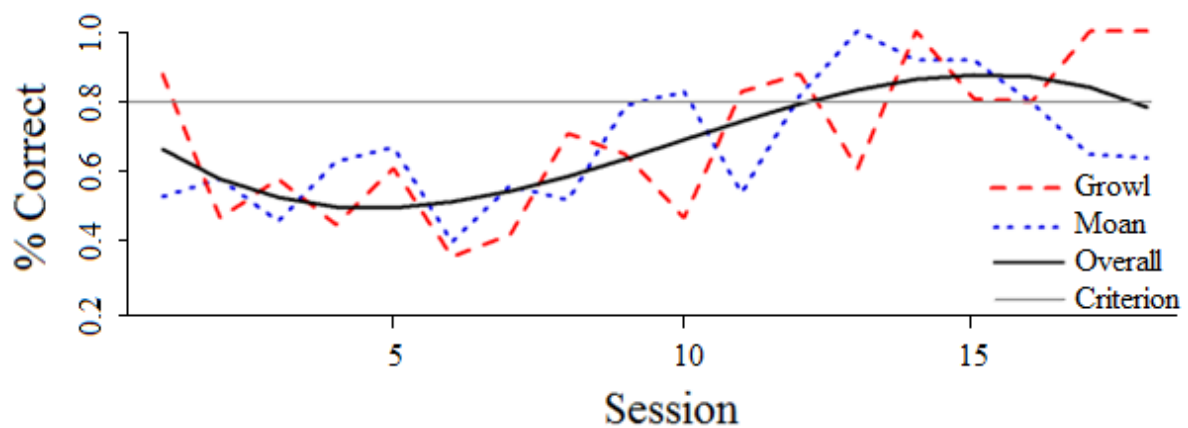
Figure 4.3. Performance accuracy by session for Zola's training period. Accuracy (percentage correct) is shown by call type (growl and moan). A nonlinear least squares model was used to determine the best fit polynomial trendline (order 3) for overall performance, weighted by the number of trials per session. Training continued until a criterion of 7 consecutive sessions with >80% accuracy for growls and moans combined was met.
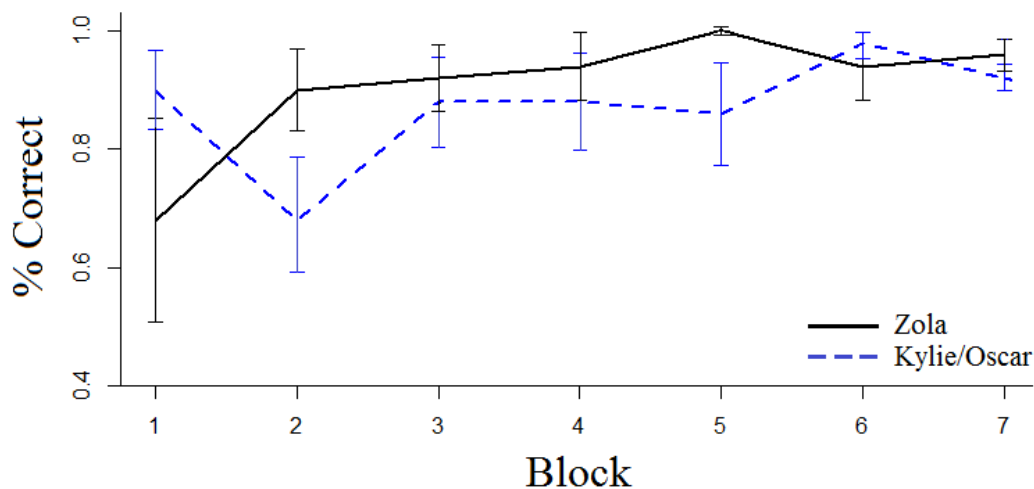
Figure 4.4. Test session accuracy by blocks of 50 trials. Accuracy (percentage correct) and standard deviation is shown, separated by test stimuli source.

than chance (χ^2 (1, N=700) = 426.05, p < 0.001). Zola's accuracy throughout testing significantly improved across sessions (Table 4.2). However, her accuracy decreased with signal exposure across sessions. Each test signal was played five times, and Zola had significantly higher accuracy during the first exposures, and lower accuracy for the last exposure. This significant decrease in accuracy across exposures occurred when hearing her own calls, but not

when hearing Kylie and Oscar's calls (Table 4.2, Figure 4.5). However, the degree to which these variables affect the model is very small. The covariates only explain a small proportion (7.5%) of the variance (Nagelkirke's pseudo r^2 = 0.075).

4.4.4   FAMD

The Kaiser-Meyer-Olkin (KMO) measure verified that sample size was adequate (KMO= 0.78). Bartlett's test of sphericity (χ^2 (903) = 89533.15, p<0.001) indicated that correlations between variables were sufficiently large. Dimensions were plotted against each other to examine clustering by call type and individual. For call type, growls and moans were separately clustered for the first three dimensions, which explained 64.16% of the data's variance (Figure 4.6a). Calls clustered by individual for the first two dimensions, which explained 53.29% (growls) and 40.90% (moans) of variance (Figure 4.6b & c).

To identify what parameters explained variance between call types, factor loadings for the first three dimensions were examined (Appendix XV). Based on the loading weights these related to noise content, frequency and duration parameters. Call clustering suggested that growls scored high on noise content and frequency but low on duration dimensions. Moans scored high on duration but low on noise content and frequency dimensions. The first dimension was also heavily weighted by categorical parameters; growls had no fundamental frequency and low signal to noise, while moans had a fundamental frequency and high signal to noise. Between individuals factor loadings on the first two dimensions were related to frequency parameters. Call clustering by individual suggested that for growls individuals differed in peak frequency parameters. For moans both fundamental and peak frequency differed between individuals, but fundamental frequency explained more of the individual variation (Figure 4.7, Appendix XV).

Table 4.2. Final generalized linear model (glm) selected using model averaging. The models used for averaging are shown in supplementary material 1. Log adjusted model weights, 95% confidence interval, and P-value are shown.

	Model Estimate e^{β}	95% Confidence Interval Lower	Higher	P-Value
Intercept	0.421	0.028	0.950	0.848
Call Source	0.736	0.220	0.965	0.381
Date	0.643	0.517	0.751	**0.026**
Signal Exposure	0.500	0.299	0.701	0.999
Date*Signal Exposure	0.470	0.447	0.493	**0.009**
Call Source*Signal Exposure	0.630	0.502	0.741	**0.047**
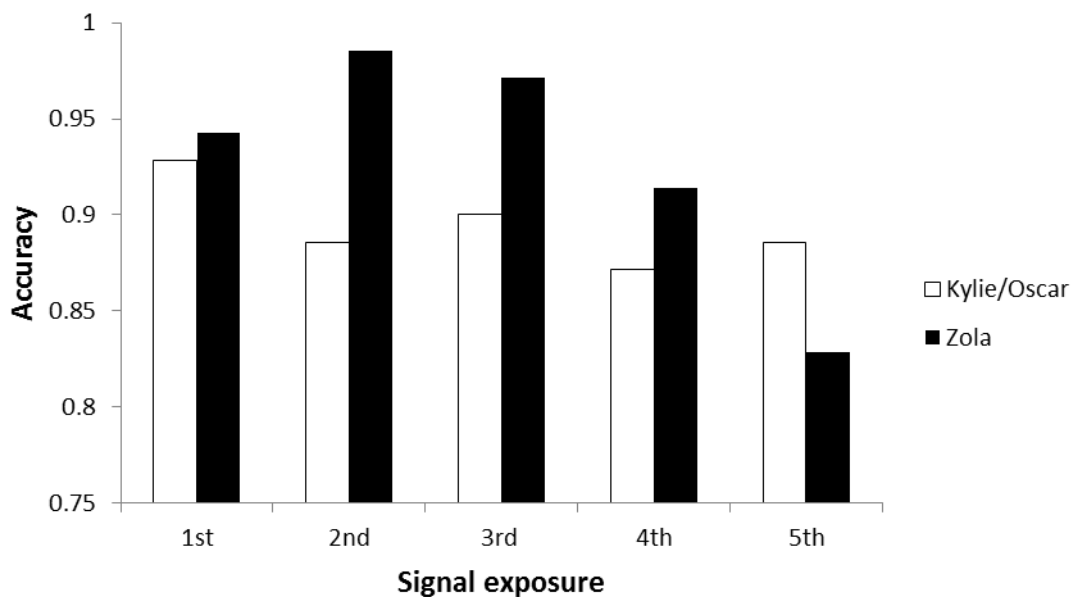
Figure 4.5 Average accuracy for the seals' performance by signal exposure. Each test signal was played five times in each test session, and accuracy for each presentation was averaged for all signals across all test sessions. As shown by the GLM model (Table 2), Zola's performance was better for the first presentations of her own calls. This was not significant for signals from the unknown seals, Kylie and Oscar.
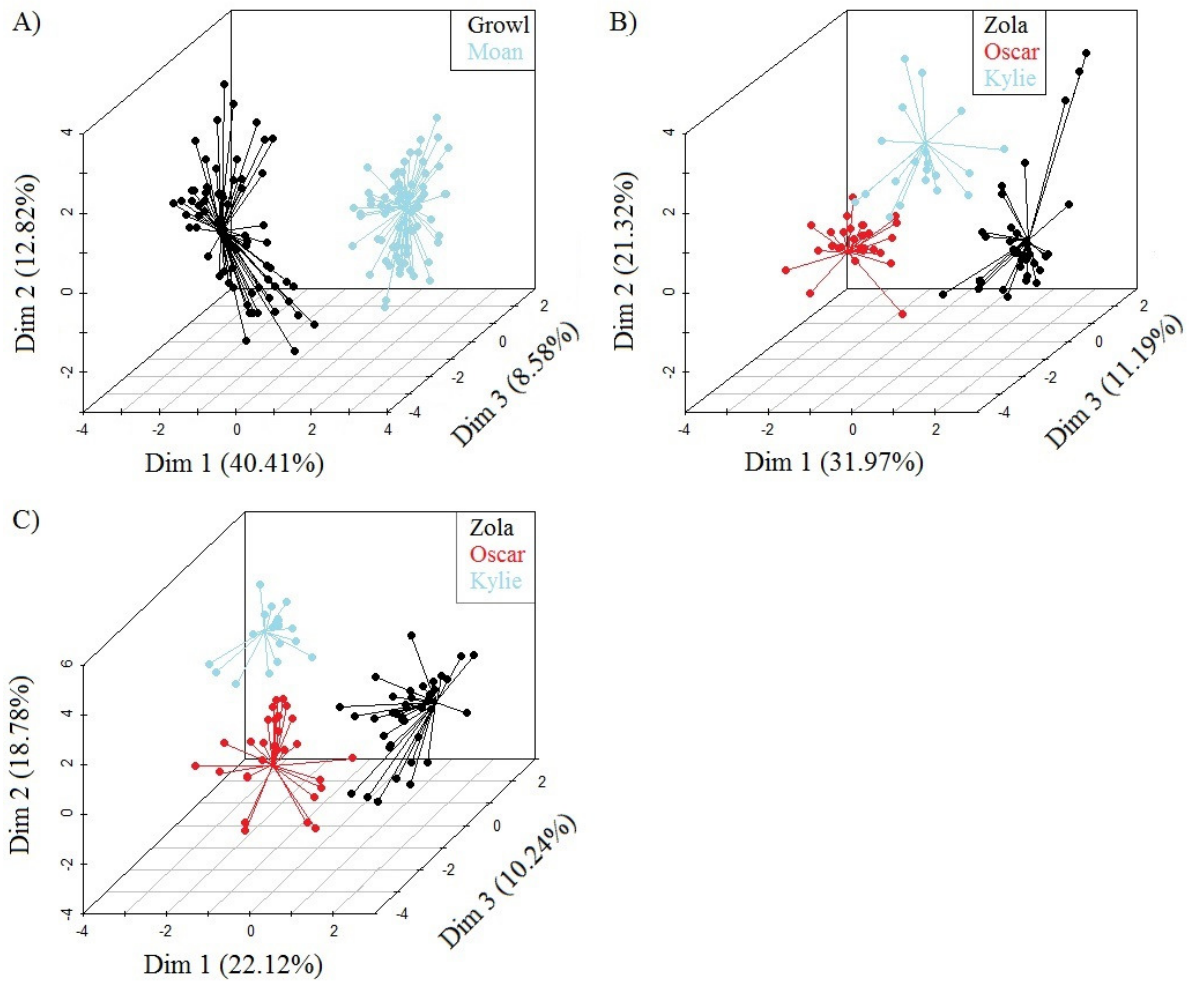
Figure 4.6. 3D scatterplots of components one through three from factor analysis for mixed data (FAMD). The proportion of the data's variance explained by each component is shown in parenthesis. A) Calls from all three seals clustered by call type. B) growls clustered by individual, C) moans clustered by individual. See Appendix XV for table of factor loadings showing how variables loaded onto the dimensions.
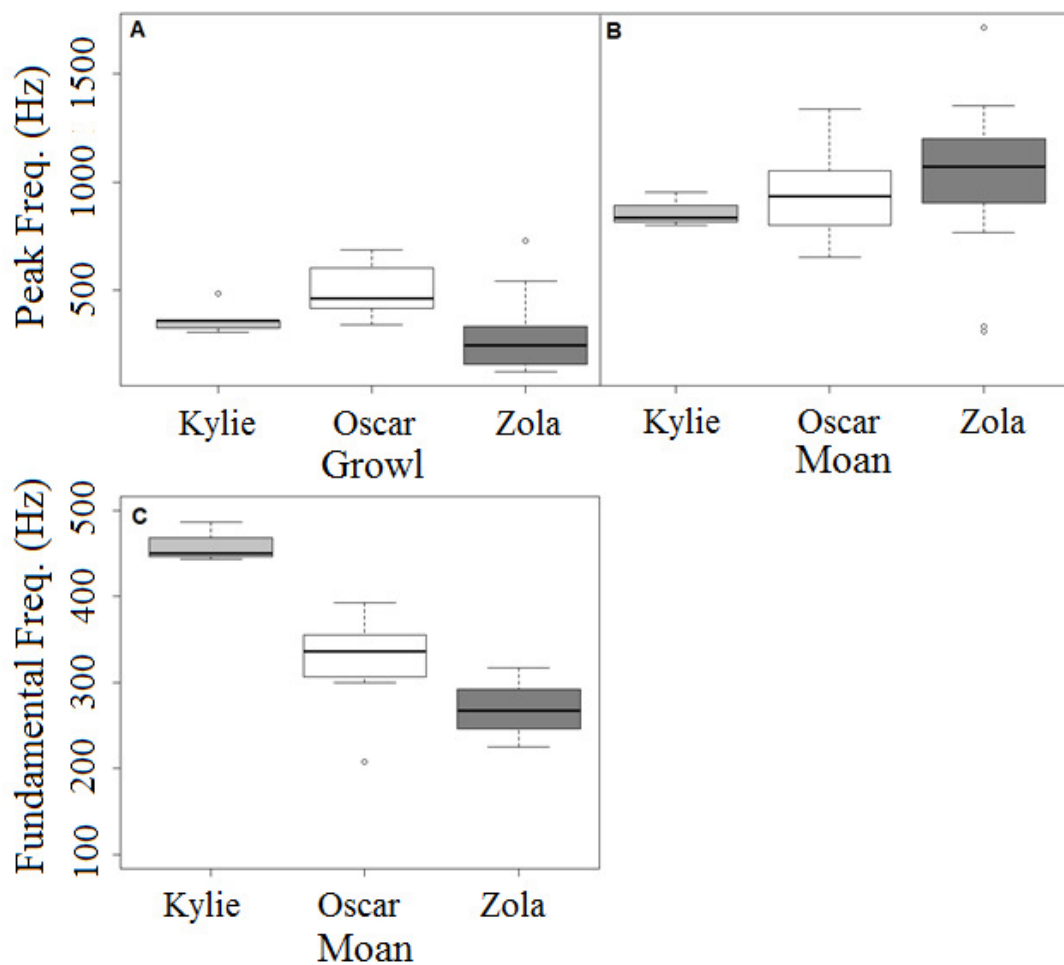
Figure 4.7. Tukey's boxplots for test stimuli by caller showing the peak frequency for A) growls and B) moans, and fundamental frequency for C) moans. Fundamental frequency is not shown for growls as they don't have a periodic structure.

4.5 Discussion

Here, I demonstrated that our seal was clearly capable of categorizing and generalizing between call classes using novel exemplars, contrasting with Shapiro et al.'s (2004) findings that grey seals were unable to generalize novel stimuli. This disparity between studies may have been due to methodological differences. Shapiro et al.'s (2004) seals were exposed to a small set of training stimuli, and then tested with novel stimuli. As the seals were trained with such a small set, they may have learned to respond to each sound as an individual cue, rather than to generalize to new stimuli. Other studies utilizing small training set have been similarly unsuccessful at training call type categorization of novel stimuli in other species (Le Prell et al., 2002; Vergara, 2011). By using

a large, continually changing stimulus set Zola was required to generalize from the start of training, and was thus successful when tested with novel stimuli.

The training and testing sets used by Shapiro et al. (2004) also varied in duration, and may have prevented the seal's successful classification of test stimuli. Zola was tested with calls of standardized duration between training and testing, which may have facilitated her success. Interestingly, although the duration of all calls fell within a relatively small range (0.7 to 1.2 seconds) to reduce variability between training and testing stimuli, call types could still be separated by duration. Additionally, Zola was two months older than the animals studied by Shapiro et al. (2004), but it seems unlikely that this would have made a difference as the animals seemed to learn the task at comparable rates.

Zola performed significantly better during the first exposures to her own calls , supporting further that she generalized signal type rather than learned items individually. The decrease in accuracy for later signal exposures may be attributed to fatigue; the later presentations of each signal would occur near the end of each block and long sessions may have impacted accuracy. Interestingly, the average duration of sessions consisting of both 50 and 100 trials were comparable (overall mean duration = 47 minutes, +/- 7). This may be attributed to variable inter-trial intervals; during some sessions, background noise was quiet and several trials were quickly completed. However, on days with varying background noise the inter-trial interval could be longer as trials did nto resume until background noise was reduced. Future studies could address this by using shorter test blocks with fixed inter-trial intervals.

The factor analysis indicated that call types could be classified using noise content, frequency, and duration while individuals could be separated using frequency. When separating call types, growls had a higher maximum frequency than moans, but lower peak frequency (Figure 4.7 and Appendix XV). Within growls, individuals were separated by peak frequency, whereas for moans although peak frequency contributed, fundamental frequency explained more of the individual variation (Figure 4.6 and Appendix XV). Thus, the seals could use different frequency parameters to provide information about both call type and caller identity.

These results are comparable to the McCulloch et al. (1999) findings that wild grey seal pup calls could be individually identified using duration and frequency parameters. Interestingly, although the calls were found to be individually distinctive, mothers did not differentially respond to playbacks of their own versus unknown pups on the Isle of May. This was not the case at a different seal colony, Sable Island, where grey seal mothers were found to respond preferentially to playback of their own pups (McCulloch et al., 1999). Although grey seal vocalisations appear to provide caller information such as for mother-pup identification, given the variable responses across populations, the function of these calls are difficult to interpret.

Interestingly, one of Shapiro et al.'s (2004) seals tended to reply to any novel stimulus with a growl. Zola similarly responded to growls by growling even in the first session of training. It is possible that the seals responded to the novel stimuli as if presented with an intruding conspecific. Later in the sessions, her growling response deteriorated, most likely due to a habituation effect in the playback context, before she successfully replied to growls by growling again after a learning period (Figure 2). The initial growling in our and Shapiro et al.'s (2004) study suggests that growls may be aggressive or defensive signals when challenged by a conspecific. This is consistent with Morton's motivational-structural code (Morton, 1977), which suggests that differences in call structure indicate emotional states within the caller such that aggressive individuals produce low frequency, noisy calls while friendly/submissive individuals produce periodic high frequency calls. While our study only examined calls produced in an artificial training context and can thus only hypothesize as to the function of these call types, future studies could examine the biological significance of these calls in wild animals using playback studies.

Given the possible biological relevance of the call types to grey seals, particularly if the growl is associated with aggressive interactions, it is even more notable that Zola was successfully able to complete the task. This demonstrates that she was able to separate motivational state not only from the production of her calls, but also her comprehension of the sound signal. I chose to use this testing method as it served as a step towards training production

learning (see chapter 5 for more information), however the inability to have accomplished the task would not have conclusively shown seals were not able to classify novel sounds signals. Instead, these results highlight how flexible pinniped vocal systems appear to be.

It is unknown how Zola used this information for discrimination. Any single parameter or combination of these parameters may have determined classification. Additionally, while parameters salient to human researchers were measured, it is also possible other unconsidered cues were used for discrimination. It is still unclear whether Zola used an exemplar based strategy, where features that best predicted group membership across multiple learned examples were used to classify novel calls, or prototype based discrimination, which is based on similarity to a 'best' example of each call type. Future research should further examine these categorization strategies by altering training stimuli systematically during playbacks.

Further studies on learning, production and perception capabilities and limitations of vocal behaviour would be valuable for pinnipeds given their interesting vocal learning skills (Janik, 2006).

Chapter 5
Trained Vocal Production Learning in Captive Grey Seals

5.1 Summary

In vocal production learning an individual's calls are modified as a result of experience with other sounds, leading to the production of sounds which are either similar or dissimilar to a model. Widely accepted evidence for vocal production learning is rare among non-human mammals. Here, I systematically tested vocal production learning in three captive juvenile grey seals. The seals were trained using positive reinforcement to vocally imitate playback stimuli. Playback stimuli were created from recordings of the seals' own calls and a human speaker which were digitally altered to vary in parameters including: number of sounds and fundamental, peak and formant frequencies. The similarity between playback stimuli and the seal's responses was evaluated by plotting call parameters and using the Euclidean distance between sounds to create dissimilarity matrices. The correlation between the signal and response matrices was evaluated using a Mantel test and compared to random chance performance. All three seals successfully matched the number of calls played per sequence. Additionally, the seals all imitated changes in frequency (either peak or formant) but did not vary fundamental frequency. These results demonstrate grey seals are capable of vocal production learning, and provide insight into the animal's vocal perception and production capabilities.

5.2 Introduction

Humans are uniquely adept at vocal communication with our capacity for spoken language. Comparatively, there are distinctive gaps in the vocal capabilities of other species. Janik and Slater (1997, 2000) separated vocal learning into two categories; contextual and production learning. Many mammalian species have demonstrated contextual vocal learning, the ability to modify comprehension and usage of calls based on experience (Janik & Slater, 1997). In comparison vocal production learning, the ability to modify call structure based on exposure to other sounds, has been documented in only a few nonhuman mammals including cetaceans (see Janik & Slater, 1997; Tyack

& Sayigh, 1997 for reviews), harbour seals (*Phoca vitulina*) (Ralls et al., 1985), elephants (Poole et al., 2005; Stoeger et al. 2012) and bats (Esser, 1994). These studies have conclusively shown mimicry of sounds outside of the species normal repertoire which cannot alternatively be explained.

However, only considering mimicry of species-atypical sounds as evidence for production learning may underestimate animal capabilities. Should vocal learners be expected to mimic such atypical signals, or is imitation of more species-typical sounds more reasonable? Past reviews have argued that production learning may be common to a much wider taxonomic distribution than previously assumed, particularly considering subtler changes in call repertoires (e.g. Janik & Slater, 1997 & 2000; Tyack, 2008). Further examination of mammalian species is needed before conclusions on the prevalence of production learning can be drawn.

While the vocal abilities of some species have been extensively investigated, further examination into comparatively less studied species is warranted. For example, vocal communication and learning in pinnipeds may be more complex than previously thought (see Schusterman, 2008; Reichmuth & Casey, 2014 for reviews). California sea lions (*Zalophus californianus,* Schusterman & Feinstein, 1965; Schusterman, 1978), Pacific walrus (*Odobenus rosmarus divergens,* Schusterman & Reichmuth, 2008), harbour seals (Ralls et al., 1985) and grey seals (*Halichoerus grypus,* Shapiro et al., 2004) are capable of vocalizing when presented with trained cues, demonstrating usage learning. Harbour seals and walruses are additionally capable of trained control over amplitude, duration, frequency, modulation and rate, and are able to produce novel calls when trained for innovation (Schusterman & Reichmuth, 2008; Schusterman, 2008). While the majority of evidence is from captive animals, in wild populations of elephant seals (*Mirounga leonine*), juvenile males preferentially produce the call types of socially dominant adult males, demonstrating usage learning and suggesting the ability for production learning (Sanvito et al., 2007).

Although the vocal abilities of pinnipeds appear to be complex, evidence for production learning is limited to one anecdotal observation. A captive

harbour seal, Hoover, spontaneously emitted vocalisations similar to human speech (Ralls et al., 1985). These vocalisations were subsequently shaped using positive reinforcement to create a repertoire of approximately a dozen English words and phrases including "come over here", and "how are you". Concurrent efforts to train another harbour seal housed with Hoover to produce speech like sounds were relatively unsuccessful (Ralls et al., 1985). Similar efforts were made to teach another male, neonatal harbour seal, "Chimo", to copy speech and physical movements by raising the pup with human models; however this was also unsuccessful (Moore, 1996).

When Hoover spoke, he was reported to sit vertically in the water with his head tucked into his neck, extended backwards, and appeared to move his tongue while keeping his mouth relatively stationary (Hiss, 1983). Human speech is produced by releasing air from the lungs through the larynx, which vibrates to generate a periodic, harmonic sound. As the sound travels some of these frequencies are emphasized or suppressed, depending on the shape and size of the mouth and pharynx. In humans, these resonant chambers can be voluntarily changed, such as by moving the tongue or changing the shape of the lips, resulting in the emphasis of specific frequencies, also known as formant frequencies (energy peaks in the sound spectrum above the fundamental). Formant frequencies are referred to by increasing frequency; formant one is the lowest frequency peak, formant two the second lowest and so forth. Humans perceive the distinction between speech sounds, such as vowels, using primarily the first two formants (though some evidence suggests the third formant provides additional information, Catford, 2001).

Hoover's ability to reproduce complex speech sounds suggests seals are capable of vocal production learning and have advanced control over the sound production system. Here, I systematically tested vocal production learning in three captive juvenile grey seals. The seals were trained using positive reinforcement to vocally imitate and mimic played stimuli, matching the number of sounds and the fundamental, peak and formant frequencies of novel sound models.

5.3 Methods

5.3.1 Subjects

Three juvenile grey seals (two female, Zola: tag numbers 73254/5, and Janice: tag numbers 73849/50, and one male Gandalf: tag numbers 73885/6) were the subjects of this study. The seals were born on the Isle of May (Firth of Forth, Scotland) in November 2011 (Zola) and 2012 (Janice and Gandalf). Post weaning (approximately three weeks old) the seals were transported to the licensed testing facility at the Sea Mammal Research Unit (St. Andrews, Scotland). The study animals were housed with other juvenile grey seals in three enclosures. All three enclosures consisted of dry, cement haul out areas and pools of varying size (one large rectangular pool (42 x 6 x 2.5 m) and two circular pools (3 x 5 x 2 m)). Training and testing of the seals occurred up to five days a week for 12 months, at which point the seals were released from captivity.

5.3.2 Acoustic Recordings

Acoustic recordings were collected using a Sennheiser MKH 416 P48 directional microphone (frequency response 40-20,000 Hz, sensitivity at 1 kHz 25 mV/Pa +/-1 dB) and Edirol FA-66 external sound card (sampling rate 96 kHz, 24-bit) onto a laptop computer. Sounds were played with the same system through an external Skytec active speaker (frequency response 32-22,000 Hz). Sounds were simultaneously played, recorded and spectrographically monitored in real time using the program Audacity 1.3 (sampling rate 96 kHz, 24-bit; Audacity Team, 2012). Weather permitting, sessions were concurrently video recorded using a Sony HDR CX250E video camera.

5.3.3 General Training Procedure

All behaviours were trained using operant conditioning and positive reinforcement. A large portion of the seals' daily diet, consisting primarily of herring (*Clupea harengus*) and sprat (*Clupea sprattus*), was used as reinforcement. Correct behaviours were reinforced with fish, while incorrect

responses resulted in a three to five second least reinforcing stimulus (LRS), with the trainer making no response, before continuing training.

All sessions took place on land and were voluntary. The seals always had access to the water, and if the animal refused to leave the water for a session, training was ended and their diet was free fed to them at the end of the day (instead of during the training or testing time). This occurred infrequently, and overall the seals were highly motivated to participate.

Captive bottlenose dolphins (*Tursiops truncatus*) have been found to develop signature whistles that match the features of the acoustic bridge (a sound such as a whistle paired with primary reinforcement, usually food) used by their trainers (Miksis et al., 2002). To limit the seal's exposure to acoustic stimuli that may impact their repertoire, a visual bridge was used at the start of training. All of the seals were initially trained for husbandry behaviours including targeting, moving between enclosures, moving in/out of the water, and stationing on land for up to fifteen seconds.

Training Procedure 2012

Training procedures differed between animals; the method for the seals tested in 2013 was modified based from experience in 2012. Initially Zola was trained with another male juvenile grey seal, Clark. Husbandry training started in January 2012. After basic husbandry behaviours were established, vocalisation behaviours were developed in May 2012. Initially the seals were reinforced for making any sound, which progressed until they were only reinforced for calling when stationed out of the water. At this point a hand cue was introduced, and the seals were reinforced for making any sound when cued in addition to staying silent when the hand cue was not present.

Once the hand cue was under stimulus control, it was paired with a sound cue. For Zola and Clark this consisted of a computer generated sinusoidal signal (0.5 seconds, 70 dB re: 20 µPa). The frequency of the signal was chosen based on the seal's mean peak frequency measured from a sample of 100 calls. For Zola, her calls had a mean frequency of 1015 Hz (SD ± 89.27), and thus a 1000 Hz signal was used for her. For Clark, his calls had a mean

frequency of 224 Hz (SD ± 34.73), and thus a 200 Hz signal was used for him. The hand cue was faded until only the sound stimulus signalled the seal's response.

At this point (June/July 2012) Clark stopped participating in any training occurring outside of the water. This appeared to be due to a problem with sea gulls, which would occasionally steal the fish used to reinforce the seals. A few occasions were noted where a gull came very close to Clark's head when hauled out of the pool. This was thought to have been aversive and resulted in decreased participation in training sessions. Additionally, Clark only vocalised with his mouth closed. This was not initially anticipated to have been a problem, and throughout training any sound produced when cued was reinforced. However, this limited the flexibility of his calls. While an attempt to continue working with Clark was made, between the lack of motivation to participate and the reduced vocal plasticity he did not proceed much farther in his training and was subsequently excluded from the study.

In June 2012 Zola was reliably vocalizing in response to the computer signal. This training signal was then digitally altered to vary in time and frequency parameters. The initial training set consisted of eight calls, which either varied in duration (0.5, 1.5 and 3.0 seconds, 1 kHz signal 70 dB re: 20 µPa spl) or frequency (unmodulated signals of 0.8, 1.0, and 1.2 kHz and 2 modulated signals which progressed from 1.2 to 0.8 and from 0.8 to 1.2 kHz, all 1.5 seconds in duration and 70 dB re: 20 µPa spl). With the longer duration sounds, the seal began to respond before the stimulus had finished. An LED light was put in eye-sight of the seal to indicate when a playback had ended. After the sound signal had finished, the light was illuminated until the seal had stopped vocalizing. Any response the seal made while the light was off was not reinforced, and within a few sessions these interruptions occurred infrequently.

The seal's response was judged by the trainer visually using a real-time spectrographic display in Audacity. The seal was differentially reinforced for producing calls that were more similar to the model. Initial reinforcement criteria were more flexible; for example when matching duration, the seal was reinforced for producing a sound within 0.5 seconds of the model. When

presented with unmodulated frequency, any sound within 100 Hz of the peak frequency was reinforced. For frequency modulated calls, any change in frequency of at least 50 Hz in the correct direction was reinforced. In trials where the spectrographic display could not be used (i.e. if the computer was too slow building a spectrogram, or if environmental noise was unexpectedly loud) the trainer subjectively determined the similarity of the match (this was estimated to have occurred in approximately 20% of trials, however no record was kept).

Zola's performance with the computer generated signals did not progress; using the set of eight computer generated training exemplars, Zola was never correct for more than a third of these trials. Thus, a new training strategy was implemented in July 2012. At this time Zola produced two distinct call types; 'growls' and 'moans'. These call types have been previously well documented; growls are noisy calls with a bandwidth of up to 20 kHz, while moans are periodic calls with a harmonic structure and bandwidth rarely exceeding 5 kHz (Figure 4.1; Shapiro et al., 2004). By September 2012, Zola was trained to discriminate between call types by matching novel growls and moans for both her own calls and those of two unknown seals, demonstrating she was capable of vocal usage learning. Further information concerning this experiment can be found in chapter 4.

Training usage learning was used as a step towards production learning; once the seal consistently matched call type, I then selectively reinforced calls that matched specific parameters. Whereas during our previous attempt I had used computer generated stimuli, I now used Zola's own previously recorded calls. As moans were most similar to the calls previously matched by the seal who mimicked human speech (Ralls et al., 1985), only moans were used from this point on. Digital manipulation of the moans was achieved using Adobe Audition 2.0 to vary the number of calls presented and the frequency of the calls presented. The frequency of calls was changed using the 'pitch shifter' function, which keeps the duration of the call constant while moving the pitch of the call. This allows the relative frequency structure of the call to remain intact while linearly shifting the frequency spectrum of the call (Lent, 1989).

Zola was trained with a set of calls with constant duration (0.5 seconds), intercall interval (0.1 seconds), and amplitude (70 dB re: 20 µPa spl) but varied in frequency and number. Peak frequency was changed in integer steps corresponding to the musical scale nearest the seal's mean peak frequency and extending more than one standard deviation based on a sample of 100 calls (mean frequency 1015 Hz, SD ± 89.27). Thus, Zola was presented with 880, 1046, and 1175 Hz calls (corresponding with musical notes A5, C6, and D6, each 0.5 seconds long and 70 dB re: 20 µPa spl). Shifting the peak frequency additionally changed the fundamental frequency of the signals (180, 210 and 245 Hz respectively). During each playback the seal was played sequences of up to three calls, consisting of every possible combination of these three frequencies. Thus, a total of 19 signals were used in this stage of training.

Previously, the seal had only been played one signal, and was immediately reinforced after responding with one call. At this point in training the seal was played multiple calls, and the seal was assumed to have finished one second after the completion of her last call. This interval was chosen as it was longer than the mean inter-call interval of a sample of 50 calls (mean = 0.27 seconds, SD ± 0.17).

Throughout this stage of training, the seal was reinforced for matching both the number of calls and the change in frequency of signals. Due to difficulties in evaluating the seal's peak frequency in real time, peak frequency was not explicitly reinforced. Instead, reinforcement was based on change in peak frequency between calls such that if the change in frequency was at least 60 Hz in the correct direction (either increased or decreased in frequency with the model), then the seal was reinforced. If the signal consisted of only one call, then the seal was reinforced for responding with only one call regardless of frequency. Testing began once the seal had reached the criterion of five consecutive sessions in which at least 80% of calls were correct matches of the model sound, which occurred in October 2012.

Testing Procedure 2012

Test trials were conducted similarly to training trials. Each trial was initiated by the seal moving into position, stationed at the side of the pool with head facing the speaker. A sound was played, and the seal was judged to have finished calling when no additional sounds were produced for more than one second. The seal's response was evaluated using a real time spectrographic display in Audacity using the same criteria as training (see above). If the seal responded correctly, she was reinforced with pieces of fish. If incorrect, a LRS of three to five seconds occured before beginning the next trial.

Test stimuli were created using one of Zolas' previously recorded vocalisations digitally altered to vary in number and frequency, similarly to training trials. While during training Zola was only exposed to up to three frequencies (880, 1046, and 1174 Hz), test stimuli extended over an octave (ranging from 698 to 2093 Hz, corresponding to musical notes F5 to C7). Throughout testing 15 'songs' consisting of different combinations of up to ten of these notes were used (Table 5.1). Each 'song' was presented with every possible number of notes (i.e. for a three note song, the first note was presented alone, the first and second note alone, and the full three note song). Thus, in total there were 92 stimuli used throughout testing.

Stimulus order was arbitrarily randomized by the researcher with one exception. In order to prevent frustration, if the seal responded incorrectly in a trial, the next stimulus presented was one the researcher subjectively thought the seal would be more likely to succeed at. This would either be a shorter sequence of the same 'song' or an alternative song the seal had higher average accuracy with ("Happy birthday" and "Twinkle twinkle", see Table 5.1). This procedure resulted in stimuli being played an unequal number of times. Every stimulus was played at least once (mean 2.924, SD ± 2.765). This average excludes the two songs which were presented when the seal responded incorrectly; they were presented the most frequently of any stimuli, with the most frequently played song ("Happy birthday") being used 30 times throughout testing. Examples of Zola's test trials are shown in Figure 5.1 and 5.2.

Song	# Notes	Musical note sequence & corresponding frequencies (Hz)
"Upscale"	8	C6 (1046), D6 (1175), E6 (1318), F6 (1396), G6 (1568), A6 (1760), B6 (1975), C7 (2093)
"Upscale skip"	4	C6 (1046), E6 (1318), G6 (1568), C7 (2093)
"Downscale"	8	C7 (2093), B6 (1975), A6 (1760), G6 (1568), F6 (1396), E6 (1318), D6 (1175), C6 (1046)
"Downscale skip"	4	C7 (2093), G6 (1568), E6 (1318),C6 (1046)
"Twinkle twinkle little star"	7	C6 (1046) C6 (1046), G6 (1568), G6 (1568), A6 (1760), A6 (1760), G6 (1568)
"Happy birthday"	6	C6 (1046), C6 (1046), D6 (1175), C6 (1046), F6 (1396), E6 (1318)
"Rocky"	6	E6 (1318) , G6 (1568), A6 (1760), A6 (1760), B6 (1975), E6 (1318)
"Les Mis"	7	A5 (880), G5 (783), F5 (698), G5 (783), A5 (880), B5 (987), C6 (1046)
"Starwars"	7	G5 (783), D6 (1175), C6 (1046), B5 (987), A5 (880), G6 (1568), D6 (1175)
"You are my sunshine"	10	D6 (1175), G6 (1568), A6 (1760), B6 (1975), B6 (1975), B6 (1975), A#6 (1865), B6 (1975), G6 (1568), G6 (1568)
"Mary had a little lamb"	7	E6 (1318), D6 (1175), C6 (1046), D6 (1175), E6 (1318), E6 (1318), E6 (1318)
"Shifted up"	4	D6 (1175), D6 (1175), A6 (1046), A6 (1046)
"Shifted down"	4	A6 (1046), A6 (1046), D6 (1175), D6 (1175)
"Extended up"	5	D6 (1175), D6 (1175), D6 (1175), A6 (1046), A6 (1046)
"Extended down"	5	A6 (1046), A6 (1046), D6 (1175), D6 (1175), D6 (1175)

Table 5.1. List of Zola's test stimuli. Zola was presented with 'songs'* which consisted of sequences of notes with varying frequency. The total number of notes and sequence (note corresponding to traditional musical scale and frequency in Hz) for each song is shown. *Note that many of the songs used are popular human tunes. These were selected as they were familiar to the trainer, which facilitated subjective judgement of the seal's accuracy for reinforcement.
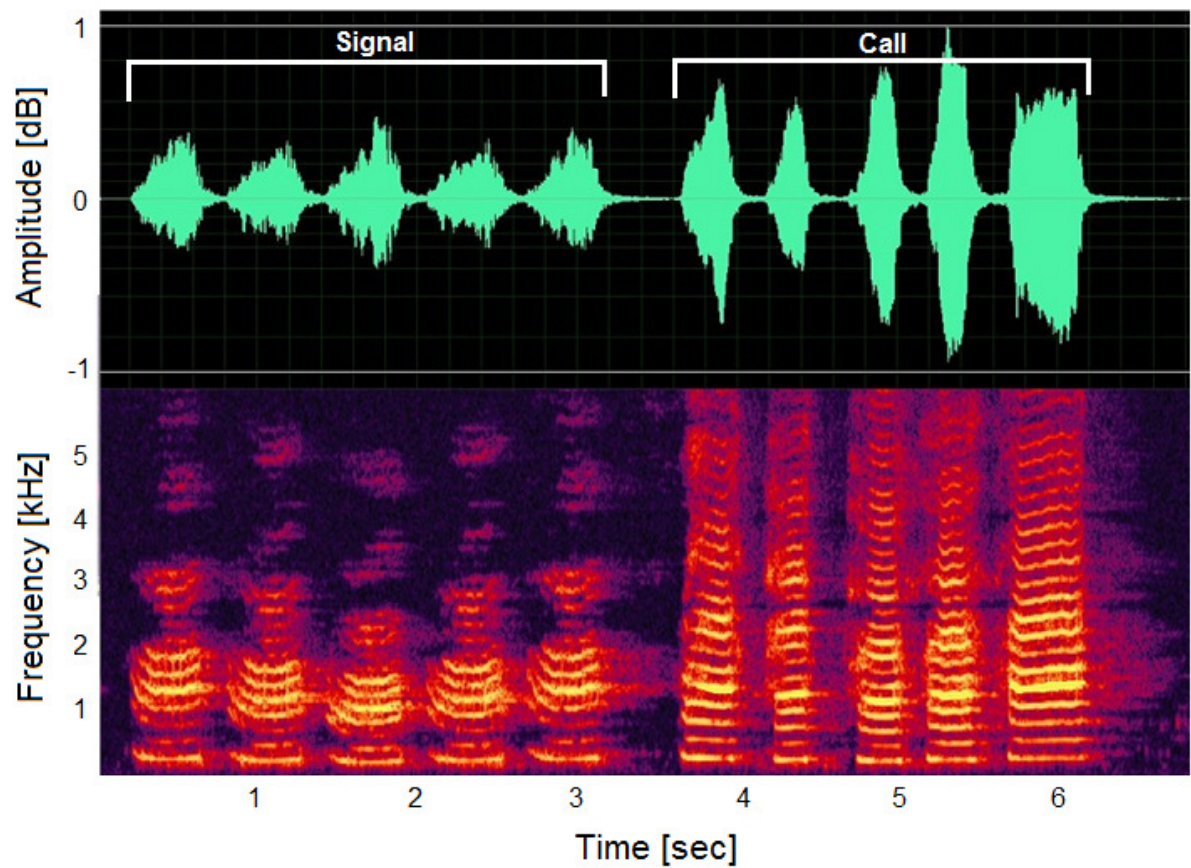
Figure 5.1. Spectrogram and waveform display of one of Zola's test trials created in Adobe Audition version 2.0 (FFT size: 2048 frequency resolution: 46.87 Hz, time resolution: 10.66 ms). The 'song' was the first five 'notes' of 'Mary had a little lamb', showing first the played signal, and then the seals response. The audio example is included in the supplemental material, listed in Appendix XVI.
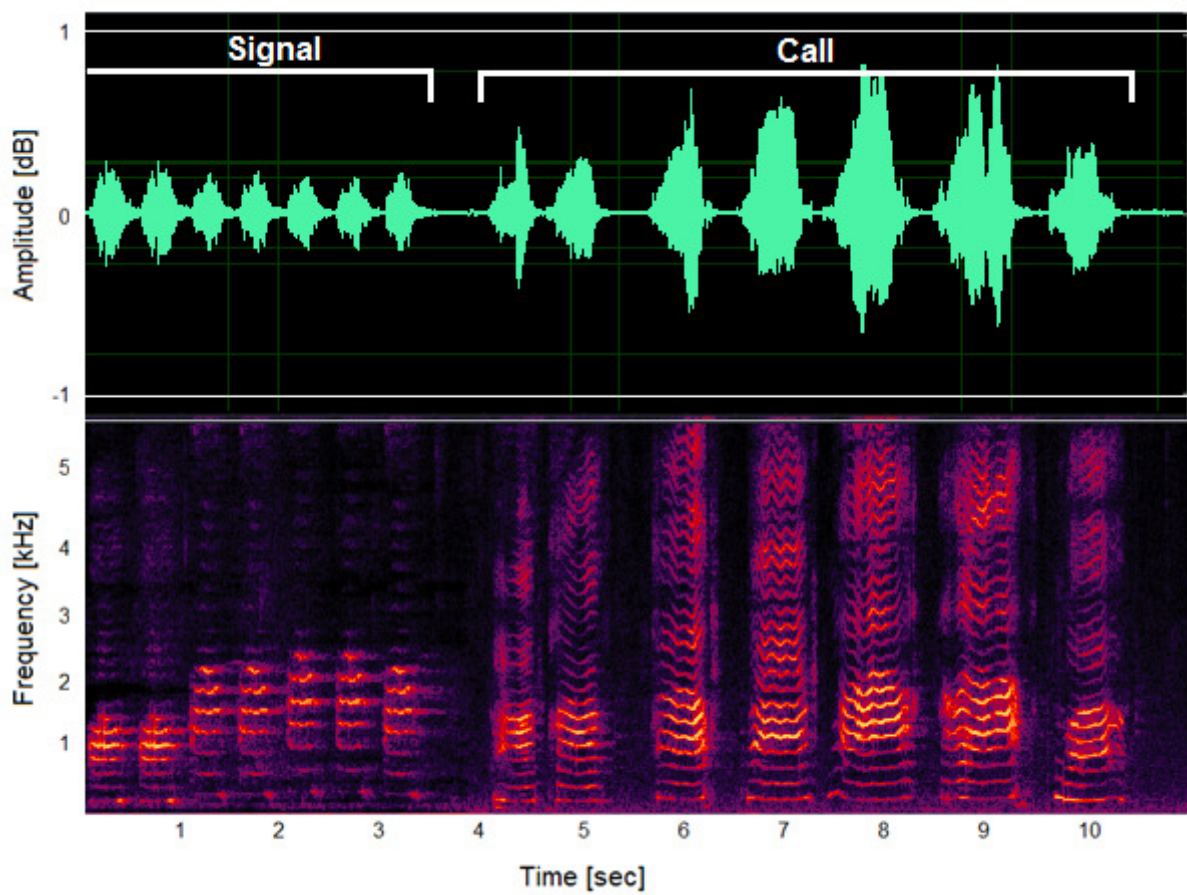
Figure 5.2. Spectrogram and waveform display of one of Zola's test trials created in Adobe Audition version 2.0 (FFT size: 2048 frequency resolution: 46.87 Hz, time resolution: 10.66 ms). The 'song' was the complete seven 'note' "twinkle twinkle little star", showing first the played signal, and then the seal's call in response. The audio example is included in the supplemental material, listed in Appendix XVI.

Training Procedure 2013

The training procedure for 2013 was similar to that from 2012, with some changes. Initial husbandry training began in January 2013 for Janice and Gandalf, and proceeded faster than in 2012. By February 2013, the seals were trained to vocalise when hand cued on land using the previously described method, but with two differences. In 2012, the seals were reinforced for making any sounds, and for Clark this resulted in only vocalizing with his mouth shut. For Janice and Gandalf calls were only reinforced if made with the mouth open.

Additionally, while in 2012 the seals stationed at a set location at the side of the pool without intentional training, both Janice and Gandalf would position themselves near the trainer, directly next to the enclosure's fence. To keep the

146

seals approximately one metre from the speaker, a physical station (a ball on which the seal positioned it's chest) was introduced. If the seal was not at station, it's responses were not reinforced.

Once the seals remained at station and the hand cue was under stimulus control, it was paired with a sound cue and then faded until only the sound cued a vocalisation response. In 2012, computer generated sounds were used with little success. Therefore, for Janice and Gandalf, recordings of the seals' own calls were used from the start. While for Zola two call types were used, in 2013 only moans between 0.3 and 0.5 seconds in duration were used. Training stimuli were composed of novel sets of ten to twenty calls, and were changed every two to three sessions. At this stage of training, the seals were reinforced for producing any single moan in response to the stimuli.

Similarly to Zola, in May 2013 both seals were trained with a new set of calls with constant duration (1.0 second) and amplitude (70 dB re: 20 µPa spl) but varied in frequency and number. Peak frequency was changed in integer steps corresponding to the musical scale nearest the seal's mean peak frequency and extending more than one standard deviation (for Janice 915 Hz, SD ± 54.08 and Gandalf 577 Hz, SD ± 31.32). Thus, for Janice 783, 987 and 1175 Hz (corresponding with musical notes G5, B5, and D6) and for Gandalf 493, 587, and 698 Hz (B4, D5 and F5) signals were presented. During each playback the seal was played between one and three calls with any combination of the three frequencies. The seals' responses were judged and reinforced similarly to 2012; the seals were rewarded for producing the correct number and change in frequency of the call. For frequency, the same cut off criterion of 60 Hz was used. In total, the seals were presented with 55 of these training signals.

The seals both had five consecutive sessions with at least 80% correct responses in August 2013, and were then presented with a new set of training stimuli. These calls had constant duration (0.6 seconds), amplitude (70 dB re: 20 µPa) and fundamental frequency. The fundamental frequency was chosen based on each seal's average for a sample of 100 calls; for Janice 380 Hz (mean 378, SD ± 33.79) and Gandalf 190 Hz (mean 192, SD ± 39.02). The calls

varied by sound spectrum (i.e. where energy was distributed across frequencies), to correspond with formants from five of the cardinal vowels; [ɑ], [e], [i], [ɔ] and [u] (Table 5.2). Vowel sounds were chosen as the harbour seal, Hoover, subjectively appeared to have more success imitating vowel sounds than consonants (Ralls et al., 1985). Additionally, previous instances of human speech imitation in an Asian elephant (*Elephas maximus*) showed higher agreement for vowels than consonants (Stoeger et al., 2012). These specific vowels were chosen as they are produced using variable mouth, lip and tongue positions (see Figure 5.3) and they were easily identifiable by human listeners. The vowels were produced by a native, North American English speaker (recorded with sampling rate 96 kHz, 24-bit) and then digitally altered in Adobe Audition to have a mean fundamental frequency similar to that of the seals. See Figure 5.4 for a spectrographic example of Gandalf's test stimuli, highlighting the formant frequencies of the vowels. The seals were presented with one sound per playback and reinforced for responding with the correct number (i.e. one call), mouth shape (subjectively evaluated by the trainer, Figure 5.3), and formant frequencies. Previously the seals' responses were primarily evaluated using a real-time spectrographic display in Audacity. However, it was more difficult to evaluate the formant energy distribution across the spectrum as quickly as necessary during training sessions. While the spectrogram was

Table 5.2. Mean formant frequency and standard deviation of Gandalf and Janice's test stimuli. Sounds were recorded from a native English speaker, and were digitally altered to match the seals average fundamental frequency. Formant frequencies were measured in Praat (settings: number of formants: 6, window length: .01, Dynamic range: 40 dB, Max formant: 6,000 Hz (Gandalf) 6,500 (Janice). Formant frequencies were measured every 25 ms, resulting in 25 measured points per call, which were averaged to obtain mean and SD.

	Gandalf				Janice			
Vowel	Mean F1	SD F1	Mean F2	SD F2	Mean F1	SD F1	Mean F2	SD F2
[ɑ]	1043.778	5.142	1571.001	39.949	1262.467	69.748	1602.133	23.913
[e]	753.667	40.5	2247.222	29.779	789.4	8.218	2692.8	10.804
[i]	333.667	74.508	2593.444	56.26	584.667	34.983	4324.867	51.964
[ɔ]	776.111	6.051	1062.336	15.427	808.667	6.821	1371.867	34.639
[u]	423.889	11.623	1073	12.684	446.133	14.535	1297.267	30.33
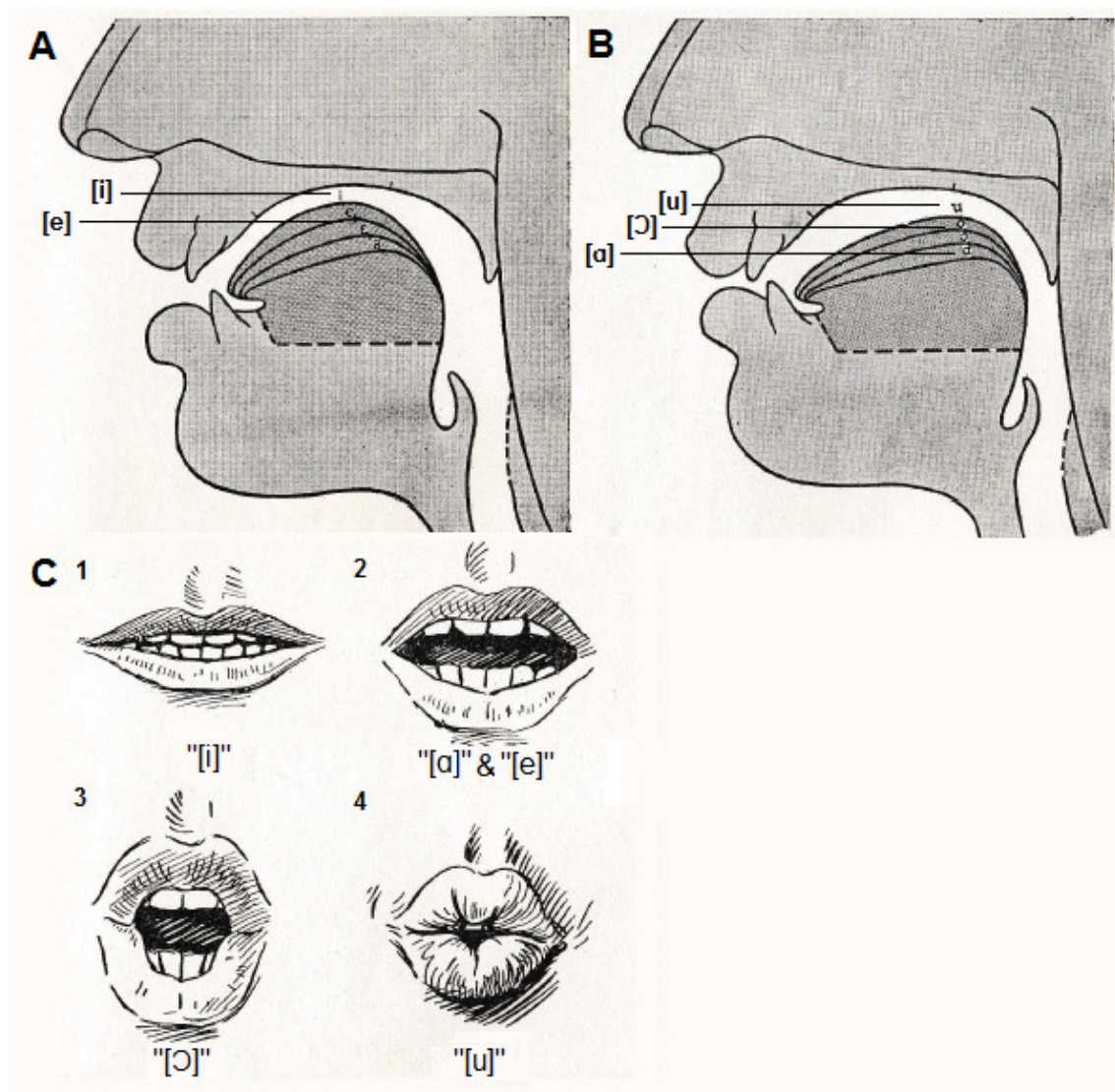
Figure 5.3. Diagrams illustrating the lip and tongue positions used to produce cardinal vowels. A) Approximate tongue positions for the 'front' vowels, in which the apex of the tongue is pushed forward in the mouth. Varying the height of the tongue in this position produces the [i] and [e] vowels. B) Approximate tongue positions for the 'back' vowels, in which the apex of the tongue is pushed back in the mouth. Varying the height of the tongue in this position produces the [u], [ɔ] and [ɑ] vowels. C) Lip positioning to produce vowels. 1) Unrounded, closed lips produce the [i], 2) unrounded, open lips produce the [ɑ] and [e], 3) rounded, open lips produce the [ɔ], and 4) rounded, closed lips produce the [u]. Images were taken from Jones (1956), and were modified to highlight the vowels used during testing.

149
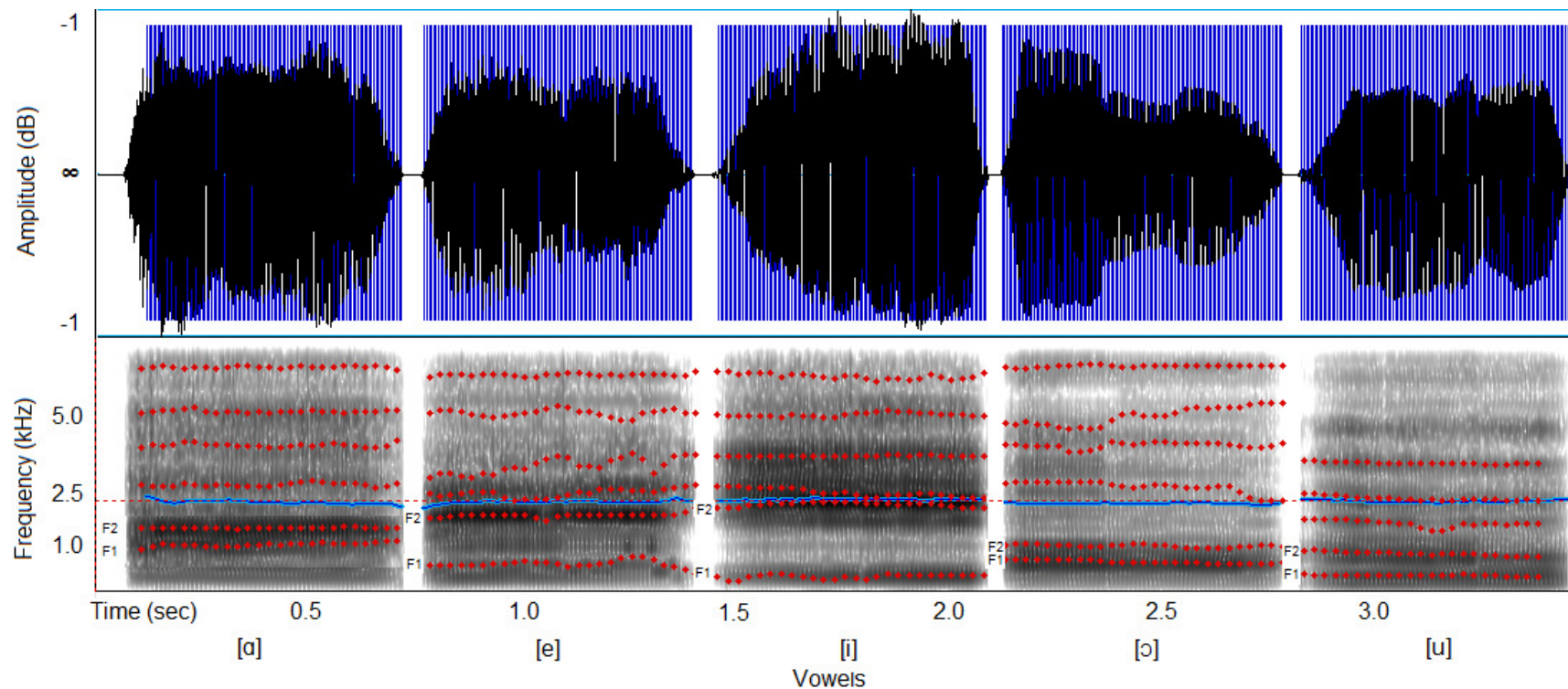
Figure 5.4. Test stimuli for Gandalf, showing the spectrogram and waveform display in Praat (number of Formants: 6, Window Length: .01, Dynamic range: 40 dB. Max formant: 6,000 Hz). Individual vowels are listed along the x-axis, with the red dots corresponding to formant frequencies. The first (F1) and second (F2) formant frequency is shown next to each vowel. The blue line corresponds to the sound's overall pitch, as calculated using the Boersma (1993) formula (approximately 2.5 kHz).

utilized to evaluate the seal's response, reinforcement was more dependent on subjective evaluation by the trainer. Once the seals had five consecutive sessions with 80% correct trials, they were then presented with multiple calls. At this time the seals were presented with up to five vowels (inter-call interval 0.05 seconds) per trial, always in the same order (i.e. vowels were always played in the following sequence: [ɑ], [e], [i], [ɔ] and [u]) in addition to being played separately. Thus, nine training signals were used at this point. Once the seals had five consecutive sessions with 80% correct responses, testing began (October, 2013).

Testing Procedure 2013

Test trials were conducted similarly to training trials. Each trial was initiated by the seal moving into position, stationed at the side of the pool with head facing the speaker. A sound was played, and the seal was judged to have finished calling when no additional sounds were produced for more than one second. The seal's response was evaluated using a real time spectrographic display in Audacity using the same criteria as training. If a seal responded correctly, it was reinforced with pieces of fish. If incorrect, a LRS of three to five seconds occured before beginning the next trial. Test stimuli consisted of the same five cardinal vowels used in training ([ɑ], [e], [i], [ɔ] and [u]). While in training these vowels were only presented in one order, during testing they were presented in randomized combinations of up to three vowels. Every possible combination was tested, resulting in 155 different stimuli, each played three times throughout testing. Thus, each seal was tested with 465 trials. An example of one of Janice's test trials is shown in Figures 5.5.

5.3.4 Analysis

The number of sounds, fundamental, and peak frequency were measured using Avisoft-Saslab Pro 5.02.04 sonogram software. Formant frequencies were measured using a software program commonly used to analyse human
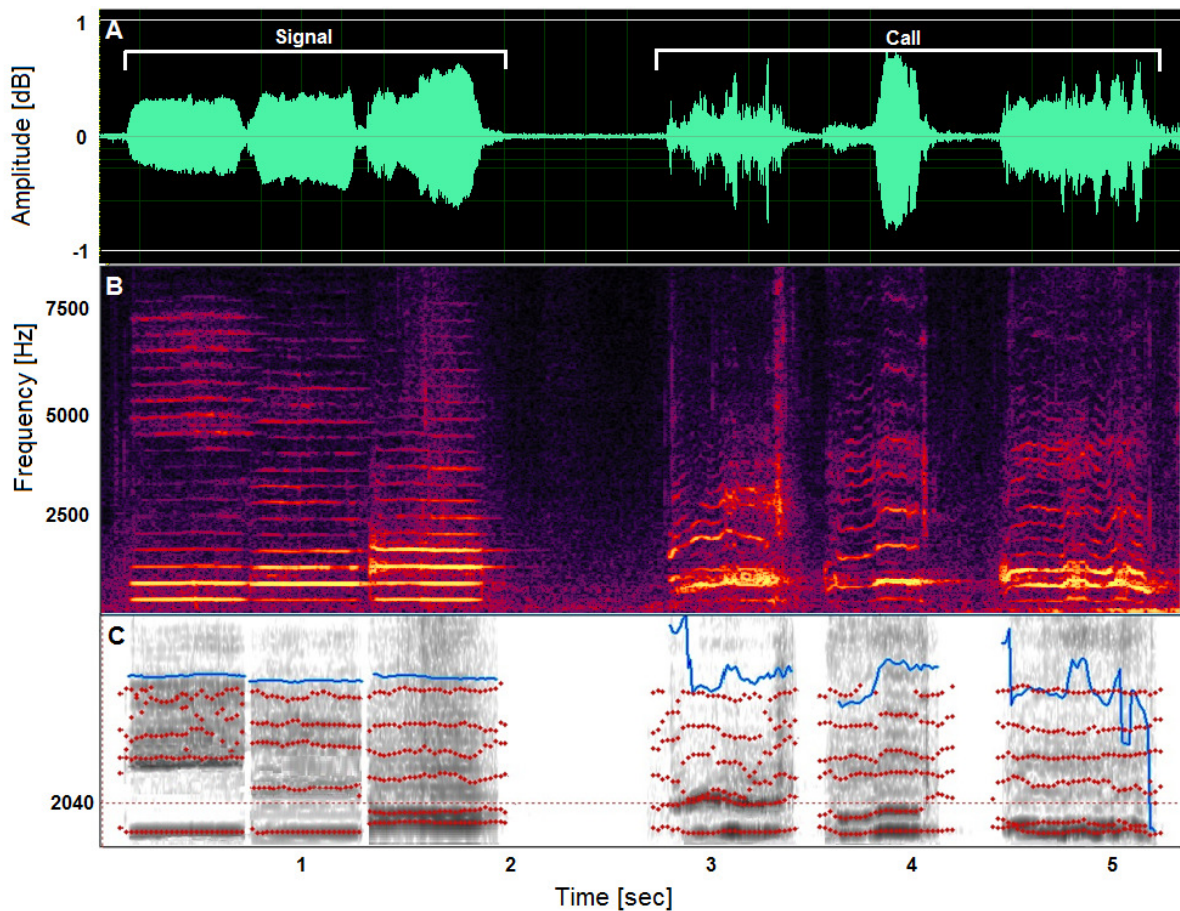
Figure 5.5. Spectrogram and waveform display of one of Janice's test trials created in Adobe Audition version 2.0 (FFT size: 2048 frequency resolution: 46.87 Hz, time resolution: 10.66 ms) and Praat (number of Formants: 6, Window Length: .01, Dynamic range: 40 dB. Max formant: 6,500 Hz). The vowels presented were [i], [e], and [ɑ], shown in the signal, and then the seal's call in response. Praat automatically calculates the formant frequencies at points throughout the call, shown by the red dots. The blue line corresponds to the sound's overall pitch, as calculated using the Boersma (1993) formula. While our seals successfully matched formant frequencies (predominantly by varying the 2nd formant), they matched poorly in comparison to previous recordings of Hoover. Audio recordings of Janice, as well as additional recordings of Gandalf and Hoover are available in Appendix XVII.

speech, Praat version 5.3.51 (Boersma & Weenink, 2014). See Table 5.3 for

definitions of the test parameters. The similarity between the signals played and

a seal's response was evaluated using distance matrices and the Mantel

statistic. As each seal was presented with different signals, statistics were run

Table 5.3. Definitions for test parameters. Number of sounds, fundamental, and peak frequency was measured using Avisoft-Saslab Pro 5.02.04 (FFT size: 1042 frequency resolution: 22 Hz, time resolution: 5.805 ms). Formant frequencies were measured using Praat (number of Formants: 6, Window Length: .01, Dynamic range: 40 dB. Max formant: 6,000 Hz (Gandalf) and 6,500 Hz (Janice)).

Parameter	Animal Tested	Definition
# of sounds	All	Number of individual sounds without a break in frequency of more than 5 ms within -35 dB of the maximum spectrum peak, ending when 1 second passed without any additional calls being made.
Fundamental Frequency	Zola	Frequency in Hz of the lowest integer multiple of amplitude peaks in a harmonic call. Measured every 5 ms and averaged across the call.
Δ Fundamental Frequency	Zola	The difference in fundamental frequency (Hz) between consecutive calls. Only measured in multiple call responses, with no measure taken for the first response in a sequence.
Peak Frequency	Zola	The frequency with the highest amplitude measured every 5 ms and averaged across the call.
Δ Peak Frequency	Zola	The difference in peak frequency (Hz) between consecutive calls. Only measured in multiple call responses, with no measure taken for the first response in a sequence.
Formant 1	Gandalf & Janice	Automatically measured in Praat as the first peak above the fundamental in the calls spectrum. Measured every 5 ms and averaged across the call.
Formant 2	Gandalf & Janice	Automatically measured in Praat as the second peak in frequency in the calls spectrum. Measured every 5 ms and averaged across the call.
Formant 3	Gandalf & Janice	Automatically measured in Praat as the third peak in frequency in the calls spectrum. Measured every 5 ms and averaged across the call.
Difference Formant 1-2	Gandalf & Janice	The difference in frequency (Hz) between the first and second formant.
Difference Formant 2-3	Gandalf & Janice	The difference in frequency (Hz) between the second and third formant.

for each animal individually. "Gower" distance was measured using the pairwise dissimilarities between sounds calculated by the daisy function in the cluster package for R 1.15.2 (Maechler et al. 2014). Daisy was chosen as it allows for distances to be calculated for multivariate mixed data. Thus, I was able to obtain a measure of overall similarity using all measured parameters in addition

to testing each parameter individually. This was done for all responses combined, as well as for a subset of the data analyzed separately. The seal's responses to the first presentation of each stimulus were analyzed separately to measure how accurate their performance was for only novel calls.

Two separate matrices were calculated; one for the signals played and one for each seal's response. Vectors between matrices were aligned such that each point matched the signal played with the seal's corresponding response. The Mantel test was then used to measure the association between the signal and response matrices with Pearson's product-moment correlation coefficient (Mantel, 1967) using the mantel function in the vegan package for R 2.0-10 (Oksanen et al., 2013). Matrix correlations were compared to chance by random reallocation of the matrix elements using 1,000 permutations, which resulted in a 95% confidence level (Manly, 1997). To avoid autocorrelation, elements were allocated such that call sequences were kept intact (Manly, 1997). This test produces two statistics, the Mantel R and p-value. The Mantel R is a value between negative one and one, where values closer to zero show no correlation, and values closer to ± 1 show a negative or postive correlation between matrices. The p-value statistic compares the correlation to chance based on the randomized permutations. Further details concerning statistical analysis are available in Appendix VXII.

One of the measured parameters was number of individual sounds in the seal's response. If the seal had been reinforced too quickly, before it would have otherwise ended its response, this might have artificially ended the seal's response and incorrectly indicated the seal was matching the number of sounds in the signal. To test if this occurred, the intercall interval was compared to the reinforcement interval (time between the end of the seal's response and delivery of fish reinforcement) for a sample of 100 calls from each seal. The mean difference between groups was compared with a Welch two sample T-test using the R stats package version 2.15.3 (R Core Team, 2013).

Throughout training, the seals were reinforced for matching playback stimuli. At the start of training, stimuli consisted of calls with a small range of frequencies (see above), but for testing, this range was extended over an

octave. If the seal successfully matched, it would be anticipated that at the start of training the seals would have produced more calls with more stereotyped frequency and by the end of training produced more calls with a wider frequency range, to match the frequencies played. To test if this occurred, the peak frequency of the first 100 calls made at the start of training was compared to the last 100 calls recorded from each seal. The variance between groups was compared with Levene's test using the package lawstat version 2.4.1 (Gastwirth et al. 2013). Levene's test was chosen because the peak frequencies were not normally distributed.

5.4   Results

Overall, all three seals successfully matched the signals significantly above chance (p≥ 0.005, Table 5.4 and 5.5). For individual test parameters, all three seals matched the number of sounds played. Zola matched the peak frequency and change in peak frequency of the signals, but did not match the fundamental or change in fundamental frequency. Gandalf matched the second and third formant, as well as the difference between formants, but did not match the first formant. Janice matched the second formant and the difference between formants, but did not match the first or third formant (Table 5.4 and 5.5).

The seals were more accurate matching some parameters than others, as shown by the Mantel R values (Table 5.4 and 5.5). All three seals were very successful matching the number of sounds (for all three animals, r > 0.92). Zola's matching of peak frequency was not exceedingly accurate (r=0.041), but was better for the change in peak frequency (r=0.151). Gandalf was not very accurate matching formant three (r=0.057) or the difference between formant two and three (r=0.043), but was better at matching formant two (r=0.226) and the difference between formant one and two (r=0.436). Janice was not very accurate matching the difference between formant two and three (r=0.025), but was better at matching formant two (r=0.336) and the difference between formant one and two (r=0.5884) (Table 5.4 and 5.5).

Overall, the seals performed comparably well for the first presentation of each stimulus (Table 5.4 and 5.5). Zola did better matching peak frequency during her first exposure. Both Gandalf and Janice did much worse matching the number of sounds on their first trials (r=0.393 and r=0.260 respectfully), but performed comparably well matching formants. The intercall interval was significantly shorter than the reinforcement interval for all three seals ($p < 0.05$, Figure 5.6). The seal's peak frequency range was smaller for the first 100 calls compared to the last 100 calls for all three animals ($p < 0.0001$, Figure 5.7).

Table 5.4. Mantel test results for Zola showing overall (all test parameters combined) and individual parameter performance. This is shown for all test responses (upper have of the table) and the first time they were presented (lower half of the table). The Mantel R score is a value between -1 and 1, where 0 shows no correlation. The p-value compares the correlation to chance based on 1,000 permutations.

ALL RESPONSES	Mantel R	P-Value
Overall	0.292	**0.000999**
# Notes	0.928	**0.000999**
Fundamental Frequency	0	0.483
Δ in Fundamental Frequency	-0.002	0.58
Peak Frequency	0.041	**0.011**
Δ in Peak Frequency	0.151	**0.000999**
FIRST PRESENTATION	Mantel R	P-Value
Overall	0.417	**0.000999**
# Notes	0.948	**0.000999**
Fundamental Frequency	0.051	0.160
Δ in Fundamental Frequency	0.067	0.154
Peak Frequency	0.115	**0.002**
Δ in Peak Frequency	0.220	**0.009**

Table 5.5. Mantel test results for Gandalf and Janice showing overall (all test parameters combined) and individual parameter performance. This is shown for all test responses (upper half of the table) and the first time they were presented (lower half of the table). The Mantel R score is a value between -1 and 1, where 0 shows no correlation. The p-value compares the correlation to chance based on 1,000 permutations.

ALL RESPONSES	Gandalf		Janice	
	Mantel R	P-Value	Mantel R	P-Value
Overall	0.247	**0.004**	0.355	**0.000999**
# Vowels	0.946	**0.000999**	0.929	**0.000999**
Formant 1	-0.016	0.845	-0.082	1
Formant 2	0.226	**0.000999**	0.336	**0.000999**
Formant 3	0.057	**0.000999**	-0.002	0.568
Difference 1 to 2	0.436	**0.000999**	0.588	**0.000999**
Difference 2 to 3	0.043	**0.000999**	0.025	**0.004**
FIRST PRESENTATION	Gandalf		Janice	
	Mantel R	P-Value	Mantel R	P-Value
Overall	0.299	**0.000999**	0.358	**0.000999**
# Vowels	0.393	**0.000999**	0.26	**0.000999**
Formant 1	-0.034	0.936	-0.103	1
Formant 2	0.276	**0.000999**	0.273	**0.000999**
Formant 3	0.052	**0.006**	0.013	0.235
Difference 1 to 2	0.468	**0.000999**	0.579	**0.000999**
Difference 2 to 3	0.104	**0.000999**	0.004	0.364
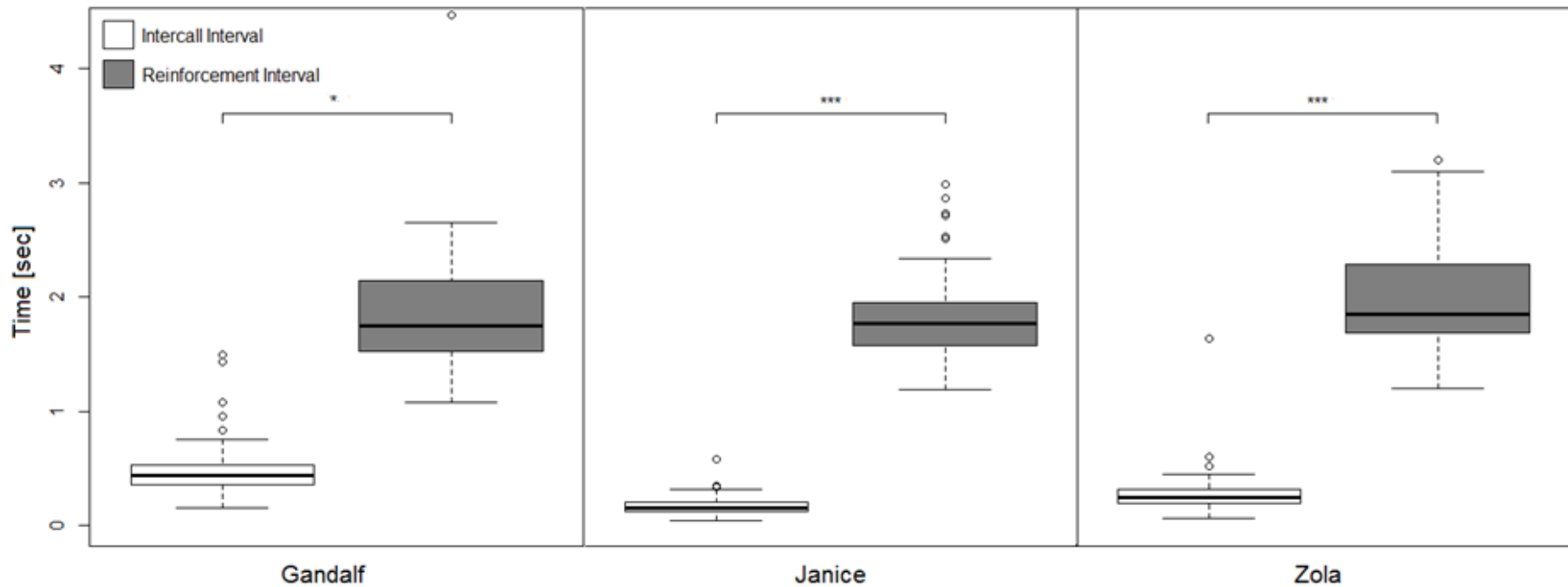
Figure 5.6. Tukey's boxplots comparing the the intercall interval to the reinforcement interval (time between the end of the seal's response and delivery of fish reinforcement) for a sample of 100 calls taken from each seal. The means significantly differed for all three seals (Welch two sample T-test for Gandalf t=-2.0887 (174.482), p = 0.038, Janice t=9.051 (119.637), p<.0001, Zola, t = 16.6219 (110.687), p-value < .0001). * indicates p<0.05, *** indicates p<0.0001.
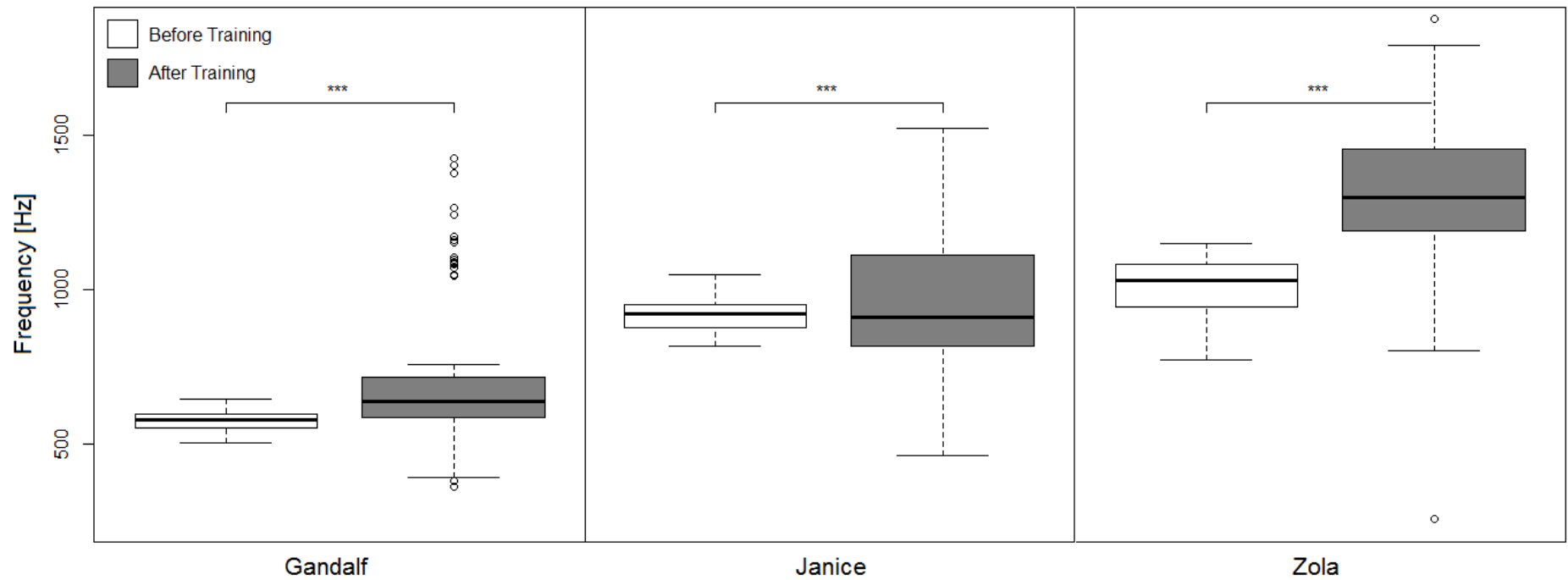
Figure 5.7. Tukey's boxplots comparing the peak frequency of the first 100 calls made at the start of training to the last 100 calls recorded from each seal. For all three seals, the last test calls had significantly greater variance than the calls at the start of training (Levene's test, for Gandalf w=93.621, p<0.0001, Janice w=95.085, p<0.0001, and Zola w= 35.397, p<0.0001), *** indicates p<0.0001.

5.5 Discussion

Here I have shown that captive juvenile grey seals are capable of trained mimicry of novel sounds. All three seals matched the number of sounds played in a sequence. This clearly shows vocal production learning, with the animals demonstrating control over sound parameters controlled by the respiratory system. The seals also showed control over their call frequencies; one seal was played sequences of musical notes, and successfully matched the model by varying the peak frequency of her calls. This was done by changing how energy was distributed across formants, rather than by shifting the fundamental frequency. The other two seals were played sequences of vowels, and successfully matched the model by varying formant frequency. Again this was done by changing how energy was distributed across the call.

The seals were more successful matching some parameters than others. All three animals were most accurate matching the number (r > 0.92 for all seals), and had lower accuracy matching the frequency of calls. This may be because matching number is presumably a simpler task (controlled by the respiratory system) than matching call frequency (controlled by the source and filter). It may also be explained by the smaller variation observed in the number of calls produced in comparison to the frequency of sounds. Sound similarity was determined using measured distances on a scale based on the variability of the calls. As the sounds frequency was more variable, this would have created more possible incorrect responses, and make an exact match less likely in comparison to matching the number of calls.

Zola poorly matched the actual peak frequency of signals, but was more accurate matching changes in peak frequency. This may partially be due to the training paradigm; reinforcement was dependent on relative change in frequency between calls regardless of absolute frequency. Gandalf and Janice both matched formant changes, predominantly by varying the frequency of the second formant. This resulted in the accurate matching of the difference between formants as well, although neither seal significantly matched the first formant and Gandalf only matched the third formant with very low accuracy (r=0.057, Table 5). By changing the position of the second formant, the

160

difference between formants changed correspondingly, resulting in a significant match even though the position of the second formant was the only one that truly varied.

Overall, the seals changed how energy was distributed across their calls to successfully match playback stimuli, as opposed to shifting their fundamental frequency. This suggests that the seals matched playback stimuli by controlling their sound's filter as opposed to the source (Janik & Slater, 2000). That is, rather than changing the tension of the vocal folds to change the sound's fundamental frequency, the seals changed their vocal tract to vary the resonant frequencies of the calls. This was unsurprising for Gandalf and Janice; the seals were presented with stimuli that differed in formants but had consistent fundamental frequencies. To successfully match stimuli, they were required to match formant changes. However, Zola was presented with stimuli that differed in both fundamental and peak frequency, but only successfully matched peak frequency. This may suggest that seals have more control over call flexibility of parameters controlled by the call's filter, as opposed to the source.

By testing the seal's ability to match varying sequences of individual sounds, I was also able to demonstrate vocal learning using biologically relevant sounds (i.e. digitally altered recordings of the seals' own calls). By combining individual sounds into new sequences, I have demonstrated that the seals were capable of "program-level" imitation. Coined by Byrne (2002), "program-level" imitation considers that few behaviour patterns are novel in themselves. Any behaviour can be broken down into sequences of muscle contractions, such that each specific movement was pre-existing in the animal's repertoire from the start. Thus, the question of novelty, and the ability to demonstrate true imitation, moves up to the level of sequences and hierarchies of movements (Byrne, 2002).

While this "program-level" imitation clearly demonstrates vocal learning, it is debatable as to if vocal usage or production learning has occurred (Janik & Slater, 2000). For Gandalf and Janice, they matched changes in formant frequencies corresponding to speech sounds which arguably do not naturally exist in grey seal repertoires. Thus, mimicry of human vowel sounds can only

be explained by vocal production learning. However, this may not be the case with Zola. Although Zola imitated call frequencies that fell outside of her normal variation, it could be argued that her calls were still within the normal repertoire of the grey seal, and thus our results could be explained by usage learning. In this case, the distinction between usage and production learning falls to Zola's perception of the stimuli; this study tested the seals' ability to match individual sounds (i.e. the different 'notes' and 'vowels') played in varying sequences. Individual sounds could have been viewed as discrete units, and the seals simply remembered the order units were presented in. If sounds were perceived as individual units which pre-existed in her repertoire and Zola simply changed how these units were combined, her performance can be explained with usage learning (Janik & Slater, 2000). Alternatively, the sequences could have been 'chunked' and viewed as one unit. If Zola combined sounds and viewed 'songs' as a single unit then each 'song' would be arguably novel and could not have pre-existed in the seal's repertoire, providing evidence for production learning.

As Zola was able to match combinations of up to ten calls, this may suggest she 'chunked' individual sounds into larger units. The working memory of humans is limited as we are only able to recall a small number (7 ± 2) of individual perceptual units (such as numbers, letters, etc.) (Miller, 1956). However, by combining these units into conceptual 'chunks', humans can remember more information (e.g. Miller, 1956). Other mammalian species are similarly limited to remembering comparatively small sets of information (fewer than 7; e.g. Kawai & Matsuzawa, 2000; Fagot & De Lillo, 2011). As Zola was able to remember sequences of up to ten 'notes', this may suggest she chunked 'notes' into larger, novel 'songs' to successfully imitate the sequences. This may have similarly occurred in Hoover; when Hoover spoke he was reported to have sounded 'inebriated' as his words slurred together (Ralls et al., 1985). This may have resulted from Hoover perceiving the strings of words as a single, chunked unit, as opposed to individual words. However, some animals appear to be capable of recalling larger sets (e.g. Inoue & Matsuzawa, 2007) and it is possible that Zola was capable of recalling the sequence order of up to ten

'notes'. It remains unclear as to if Zola's matching demonstrates usage or production learning.

At the start of training the seals' calls were more stereotyped, with peak frequency falling within a narrower range, compared to the calls produced at the end of testing. This is consistent with the training paradigm; the seals were reinforced for matching frequency parameters, which had a greater range progressing from training to testing. While this change is most likely due to the training paradigm, the seals studied were juveniles and call changes could be attributed to physical maturation. However, this is unlikely as results from chapter 2 indicated that seal calls become less variable with age. The extended peak frequency range can be most likely attributed to learning and the training paradigm. As previously discussed, this also suggests that perhaps Zola was capable of vocal production learning; these results clearly show that at the start of training Zola only produced calls within a narrow range of peak frequencies, and by the end of the testing calls extended past her previous repertoire range. However, just because Zola did not previously produce these calls during training does not necessarily indicate they were not a part of her repertoire. Through positive reinforcement training, these calls may have simply become more likely to be produced during testing such as with 'action-based learning' (Marler & Nelson, 1992). In songbirds, 'action-based learning' occurs when during the plastic phase of development individuals overproduce song, and during crystallization specific songs are selectively socially reinforced, leading to certain songs being reproduced while others are discarded (Marler & Nelson, 1992). My training paradigm may have similarly modified Zola's repertoire.

Throughout testing, care was taken to avoid any inadvertent cueing of the animal. This was particularly a concern for matching the number of sounds. Animals are capable of using subtle cues to answer correctly during simple tasks such as numerical counting, as made infamous by the horse 'Clever Hans' who correctly answered arithmetic questions using variations in human body position (Pfungst, 1911). Although trainers were kept out of the animal's sight, the delivery of reinforcement could have inadvertently cued the seal to end its response. However, care was taken to prevent this by ensuring the

reinforcement interval was significantly longer than the intercall interval. An LED light was used to indicate when played signals were finished, cueing the seals' responses. Potentially the light turning off could also have cued the seal to stop vocalizing. However, the LED light was turned off at the same time reinforcement was delivered. It is thus unlikely that the researchers inadvertently cued the animals. Since the experimenter was not visible to the seal, frequency most likely could not have been inadvertently cued.

Given the advanced vocal production learning abilities of humans there are distinctive gaps in the comparative vocal capabilities of other species, particularly in our evolutionarily closest relatives, nonhuman primates. While the paucity of evidence across nonhuman mammals may indicate distinct differences in the vocal abilities between species, alternatively it may be attributed to how we test production learning abilities. All of the currently accepted vocal learning species have displayed arguably advanced control over their call frequency parameters. Matching frequency parameters is a presumably complex capability entailing advanced neurological, perceptual and production control. The seals were successful matching such advanced frequency changes, but other species may not be similarly capable. Limited motor control over the vocal structures is one potential hypothesis for why nonhuman primates are incapable of vocal production learning (Fitch, 2000). Production learning may be more widespread for call parameters which are simpler to control, such as number, amplitude and duration of calls.

Janik and Slater (1997 & 2000) discuss how equally compelling evidence for production learning comes from control over a less complicated aspect of the vocal tract, the respiratory system. Although the currently accepted non-vocal learning species may not be able to control the call frequency, they may be capable of manipulating sound parameters controlled by the respiratory system such as amplitude, duration and call rate, which are often not considered in most of the vocal production literature (Janik & Slater, 1997, 2000). Here, all three seals matched the number of calls produced with high accuracy. Other species appear to be similarly capable; for example an orangutan (*Pongo* spp), "Bonnie", was observed to spontaneously produce

164

whistles similar to a human model (Wich et al., 2009). With minimal training, Bonnie matched number (one versus two) and duration ('long', mean = 2.3 sec, SD ± 0.5, and 'short', mean = 0.5 sec, SD ± 0.1) of a human model (Wich et al., 2009). Although whistling is not truly a vocalisation, it is controlled using the lips and suggests that perhaps non-human primates may have some capacity for production learning. Andrew (1998) proposed that the advance control primates have over mouth movements, such as shown by lip smacking, is an evolutionary precursor to language development. This is particular interesting, given that speech is distinguished by formant differences that are partially controlled by the lips (Janik & Slater, 2000). This has also been shown in other species; bottlenose dolphins are capable of mimicking human speech by matching the number and length of syllables (Lilly, 1965). Similarly, "Noc", a beluga whale (*Delphinapterus leucas*), was reported to spontaneously produce sounds structured similarly to human speech in duration and number of syllables (Ridgway et al., 2012).

As such, examining sound parameters other than frequency may be more relevant for some species, and may reveal that a wider range of species are capable of vocal production learning. The currently accepted black or white classification of species as either vocal learners or non-vocal learners underestimates production learning abilities. Vocal learning is better measured using a continuum, considering more advanced abilities to control the sound's frequency separately from parameters controlled by the respiratory system.

Mimicry of speech sounds in nonhuman animals is potentially difficult due to physiological differences. Fundamental frequency is predominantly determined by the length and tension of the vocal cords, while formant frequencies are determined by the vocal tract length and shape of the resonance cavities (Fitch & Reby, 2001; Fitch, 2006). The position of the human larynx is fairly low in comparison to most mammals, which presumably allows for greater flexibility in the resonance tract to create variable sounds, such as speech (e.g. Fitch, 2000; Fitch, 2005). Other mammals appear to alter body positions to mimic human speech. For example, an Asian elephant, Koshik, mimicked Korean words by using his trunk to manipulate the structure of his

oral cavity (Stoeger et al., 2012). When the harbour seal, Hoover, spoke, he was reported to be positioned vertically in the water, with his neck retracted and head pushed backwards (Ralls et al., 1985). While Hoover's mouth remained relatively stationary, his tongue appeared to move position (Hiss, 1983). In my test paradigm, the seals were required to sit on land, with their head facing the speaker. Although the seals significantly matched test parameters, they did not produce calls that were as similar in comparison to Hoover's performance (see Appendix XVIII for an example of Hoover speaking). This could potentially be attributed to requiring the animals to station in a fixed position. The seals may have shown more flexibility if allowed to move the head and neck more freely. Future investigations could allow more flexibility in body position during testing to potentially encourage better matches.

Our understanding of vocal production learning is predominantly based on extensive study of songbirds, which may provide some insight into the capabilities of mammals. In many songbirds there is a developmental 'sensitive period' for learning, during which time an individual is receptive to a tutors' songs. This developmental window varies between species, and typically individuals do not start to reproduce tutored songs until testosterone levels increase during sexual maturation (Catchpole & Slater, 1995). Vocal learning in mammals may follow a similar pattern. The harbour seal, Hoover, was orphaned as a pup and raised by a Boston fisherman, Mr. Swallow, for four months before moving to the New England Aquarium (Hiss, 1983). Hoover was first reported to mimic sounds five years later. The words that Hoover mimicked were very similar to phrases that Mr. Swallow regularly repeated to Hoover, such as "come over here" and "how are you". Reportedly, Hoover sounded very similar to Mr. Swallow, speaking with a New England accent, such as "Come ovuh heah" and "how wah yah" (Hiss, 1983). Hoover appeared to learn similarly to songbirds; he was exposed to the sounds as a pup, potentially during a 'sensitive period' of development, and upon reaching sexual maturity (between three to seven years in male harbour seals; Atkinson, 1997) began speaking. Hoover's speech was remarkably accurate at this time, comparably to some songbirds which reproduce tutor songs with very little practice. Koshik the Asian

elephant was also first noticed producing speech-like sounds near sexual maturity (Stoeger et al., 2012). Our seals matched calls as juveniles, suggesting that production learning is not entirely restricted by physiological age limitations. However, the seals may produce better matches upon reaching sexual maturity. Future recordings of our seals would be valuable to investigate if they incorporated playback stimuli into their natural repertoires, and if the quality of the matches improves upon reaching sexual maturity.

In some songbirds, social experience also impacts upon vocal learning. Learning from live tutors is more effective than from taped recordings, and can extend the length of the sensitive period (e.g. Baptista & Petrinovich, 1986; Nordby et al., 2001). Training paradigms may similarly affect song learning. Zebra finches are unable to learn songs from taped tutors (Eales, 1989), but are able to do so when given operant control over the taped song in a training paradigm (Adret, 1993). Social interaction and training may similarly impact vocal learning in mammals. When Hoover was adopted as a pup, he was hand fed and exposed to regular human contact. The phrases he learned were frequently said when his caregiver initiated social interaction, such as when Mr. Swallow came home and called out to Hoover (Hiss, 1983). Koshik was raised with humans from birth (Stoeger et al., 2012). For seven years, Koshik was exclusively exposed to humans (Stoeger et al., 2012). This exclusive social interaction with human models during early development may have contributed to Hoover and Koshik's spontaneous speech mimicry. With our seals, the training paradigm may have facilitated social interactions and encouraged learning which would have otherwise not occurred naturally in juvenile seals.

In songbirds, song is primarily attributed to being produced by males, potentially due to song production being initiated by elevated levels of testosterone (Catchpole & Slater, 1995). However, females of many species have also been observed to produce song (e.g. Odom et al., 2013), particularly when they have elevated testosterone levels (Garamszegi et al., 2007). This suggests that while both sexes are capable of song learning, production of learned song is linked to hormonal states. However, vocal production learning in mammals does not appear to be sex linked (e.g. Jones & Ransome, 1993;

Miksis et al., 2002; Foote et al., 2006). Here, I demonstrated vocal production learning for both male and female grey seals. While in natural contexts there could be a sexually dimorphic predisposition for call learning, the use of training in a captive environment suggests both sexes are capable of production learning.

Vocal production learning is a complex skill, and is presumably costly. Having such cost suggest there would be substantial benefit to individuals using the ability. Imitation, in which individuals copy the sounds of conspecifics, is primarily thought to facilitate sexual selection and territorial aggression in songbirds (Catchpole & Slater, 1995). This may be similar in mammals; juvenile male elephant seals and killer whales match the call types of adult males (Sanvito et al., 2007; Crance et al., 2014), preferentially those of dominant individuals (Sanvito et al., 2007). Vocal learning could also foster social relationships; in dolphins individuals occasionally copy or match the signature whistles of other group members, usually those they are in close social relationships with such as mother-calf pairs and male alliance partners (Janik & Slater, 1998; Janik, 2000; King et al. 2013). This suggests that whistle matching may be an affiliative signal for maintaining social bonds (King et al., 2013). Less is known about the function of mimicry, in which individuals change their calls to be more similar to models other than conspecifics. In our study, I had little success training the seals with computer generated sounds, but was more successful training the animals with altered playbacks of their own calls. Although human vowels were used, they were altered to be more similar in frequency parameters to the seals' own calls.

Our seals were capable of both imitation and mimicry, however using sounds more similar to the animals normal repertoire appeared to facilitate learning. This is particularly interesting when considering if usage or production learning occurred with Zola; Zola was presented with novel sequences of musical notes. If she had been able to match a synthetic tone, this would have provided conclusive evidence for production learning. However, as she was only able to match using altered versions of her own call, thus these sounds

may have pre-existed within her repertoire and matching can be explained by usage learning.

We are only beginning to understand the complex vocal capabilities of pinnipeds, which appear to be relatively flexible in learning, production and perception capacities. This is especially interesting given that many phocid seals lack rigid social structures, predominantly spending time isolated at sea with the exception of breeding and pupping seasons during which time dense groups of seals congregate (Pomeroy et al., 2005; Ruddell et al., 2007). Perhaps vocal production learning is favored in species with such unstable social structures, given that the most likely function of vocal learning is to facilitate social interactions. Without rigid social structures, learned vocal signals may be more important for mediating social interactions. This is also interesting given that many phocid species produce complex vocal sequences during mating season which fit Thorpe's (1964) definition of bird song (e.g. Green & Burton, 1988b; Morrice et al., 1994). Perhaps vocal production learning in pinnipeds plays a role in mate selection comparably to bird song.

Despite their apparently advanced capabilities, we currently lack a basic understanding of the natural repertoire of many pinniped species, their vocal development, and the limitations of their perceptual and production capacities for learning. Further investigation of the vocal abilities of pinnipeds, both in wild and captive populations, would be valuable as they appear to be ideal models for mammalian vocal learning.

Chapter 6
Comprehension Learning in Grey Seals:
Use of Acoustic Tags as Signals to Find Fish*

6.1 Summary

With vocal usage learning, animals learn to produce calls in specific contexts. This involves the animal acting as a signaller, and their calls provide information to others. Conversely, vocal comprehension learning allows for individuals to act as the receiver, and interpret the meaning of perceived calls from others as a result of experience. Comprehension learning is not limited to conspecific calls; anthropogenic noise is often introduced systematically and could provide information, such as for prey detection. Here, I show that grey seals spontaneously learn to use sounds from acoustic fish tags as an indicator of food location. In 20 randomised trials, 10 grey seals individually explored 20 foraging boxes, with one box containing a tagged fish, one containing an untagged fish and all other boxes being empty. The tagged box was found after significantly fewer non-tag box visits across trials, and seals revisited boxes containing the tag most often. The time and number of boxes needed to find fish decreased throughout consecutive trials. Two additional tests were conducted to further investigate the role of the acoustic signal: 1) tags were placed in one box, with no fish present in any boxes and 2) additional pieces of fish, inaccessible to the seal, were placed in the previously empty 18 boxes, making possible alternative chemosensory cues less reliable. Here, the acoustically tagged box was found faster than the control box. These results show that seals learn to use information from anthropogenic signals to locate food.

6.2 Introduction

Animal communication relies on individuals either acting as a signaller, providing information to others, or a receiver, interpreting provided information. Previously I discussed vocal usage learning, where animals learn to produce calls in specific contexts as a result from experience. This involves the animal acting as a signaller, and their calls provide information to others. Conversely, vocal comprehension learning allows for individuals to act as the receiver and interpret the meaning of these biological sound signals, as learned through experience.

Anthropogenic noise is often systematically introduced, and can be beneficial to some species depending on context and how others react to it. Rather than vocal comprehension learning, as the sounds are not biological, these cases illustrate acoustic comprehension learning, the learned contextual significance of any sound signal. There are several ways animals can exploit increased noise levels; masking by anthropogenic noise can protect prey from acoustic detection by predators (Barber et al., 2010) or conversely increase foraging success of predators by preventing acoustic detection by prey (Chan et al., 2010). Western scrub jays (*Aphelocoma californica*) prey upon eggs of nesting species, but avoid areas with increased noise (Francis et al., 2009). In this case, noise pollution decreases nest predation and thereby increases reproductive success of the prey species (Francis et al., 2009). Such benefits of noise may explain the increased success of some birds in habitats with extensive human activity (Slabbekoorn & den Boer-Visser, 2006; Clergeau et al., 2006).

The use of sound from a localised acoustic source can also facilitate learning by indicating a location of interest. Acoustic deterrent devices (ADDs) aim to elicit avoidance responses in aquatic predators, such as seals, and are currently being used to reduce depredation in fisheries. However, seals that have previously found fish at a location close to an ADD quickly habituate to these sounds (Götz & Janik, 2010; Götz & Janik, 2013). Observational evidence suggests that ADDs may also attract predators (Bordino et al., 2002) and in such cases may even cause higher incidences of predation (Jefferson & Curry,

1996) due to conditioned associations between sound and prey, also known as the 'dinner bell' effect. If anthropogenic noise regularly serves as such a signal, it may influence animal behaviour and ecology much more widely than previously assumed.

It is possible that another acoustic device may attract seals to prey. Acoustic coded transmitters, also known as pingers, are currently being widely used to monitor marine fish and invertebrate species (Cooke et al., 2011). The pingers are placed within the fish's body cavity or attached dermally. At specific intervals the device emits an acoustic signal which provides information about the animal's movement and environment. Typically ultrasonic frequencies are used as they are not perceptible by the fish and invertebrate species studied. As the target animal is not able to hear the emitted signal, it is assumed the tag does not alter natural behaviour.

While the target species is not sensitive to the pinger, other animals in the environment, including predators, may potentially detect it. If a predator perceived the acoustic signal, it is possible that the sound would be associated with the presence of prey. Such tracking devices could inadvertently cause a higher incidence of predation on fish tagged with pingers through a learned 'dinner bell' effect.

The majority of pingers currently in use produce signals ranging from 50 to 100 kHz at 130 to 160 dB re: 1 µPa (Bowles et al., 2010). Although the tagged fish species are not sensitive to the acoustic signal, it is audible to some predators, including seals (Bowles et al., 2010; Cunningham et al., 2014). Sensitivity thresholds for a reproduced Vemco 69 kHz fish tag were measured and used to estimate detection distances for a harbour seal (*Phoca vitulina*) and California sea lion (*Zalophus californianus*) (Cunningham et al., 2014). Both species were capable of detecting a 69 kHz Vemco coded tag signal (source level of 165 dB re: 1 µPa) at simulated distances of greater than 200 meters (Cunningham et al., 2014).

While this past work has shown that seals are capable of perceiving fish tag signals, it is unknown whether they learn to utilize this information for prey

detection. Here, I examine if captive grey seals (*Halichoerus grypus*) learn to use sounds from fish tags as an 'acoustic beacon' to find food.

6.3   Methods

6.3.1 Subjects

Ten juvenile grey seals (three females, seven males), born on the Isle of May (Firth of Forth, Scotland), were the subjects of this study. Four of the seals were born in November, 2011 and six in November, 2012. The seals had been followed from birth, had never been in the sea, and had no previous experience associating anthropogenic sounds with food. After weaning (at approximately three weeks of age) pups were transferred to the licensed captive facility at the Sea Mammal Research Unit (St. Andrews, Scotland). The seals were fed a varied diet of several fish species, however during testing only whole adult herring (*Clupea harengus,* approximately 100g in size) was used. Throughout testing, all of the seals were between three and eight months old. They were all released back into the wild within one year of initial capture.

6.3.2   Testing Enclosure

The seals were tested in a 37.5 x 6 x 2.25 m concrete pool. Twenty foraging locations were equally distributed around the bottom of the pool (Figure 6.1). Each foraging spot consisted of a PVC pole suspended from the surface with a chamber at the base. The chamber consisted of a 35 x 30 x 40 cm box, a 25 x 34 cm bucket and a 14 x 16 cm door flap (Figure 6.2). The fish were hidden inside the box; to retrieve a fish, the seals put their heads into the bucket through a door flap. The bucket allowed the seals head to enter the box, but restricted how far into the box the seal could reach. Fish was placed either on the plate, where the seal could reach them, or below the plate, where the seal could not access the fish. From within the box, the fish could then be taken from the plate where it was secured with an elastic band. Magnetic reed switches on the door flap and the plate holding the fish interfaced with a customized Matlab® program which logged door-opening and fish removal events. The program recorded the location and time of each box visit and/or fish retrieval.

Figure 6.1. Photograph of the testing enclosure, drained of water. Twenty foraging locations were distributed around the pool. At each location the seal could place its head through a bucket to access fish hidden within the box.
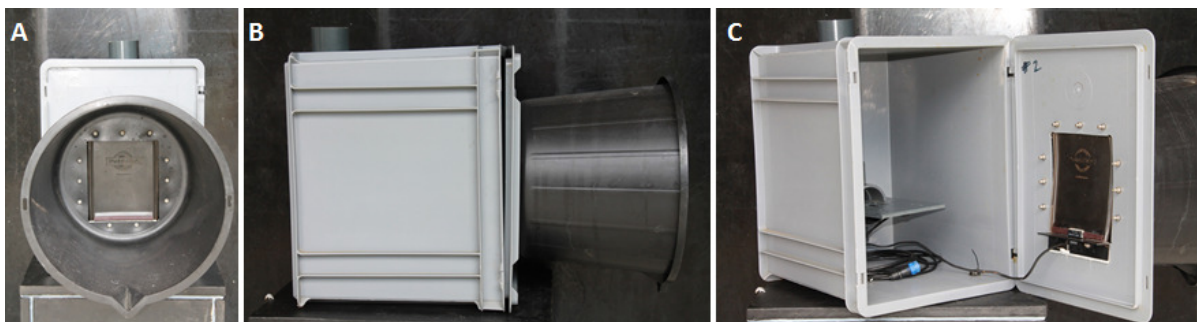
Figure 6.2. Photograph of the foraging boxes showing the view from the A) front of the box, B) side of the box and C) inside of the box. To access food, seals placed their head into the bucket, and through a door flap to take fish secured on a plate with elastic bands. Time when the door flap opened or closed and when the fish was taken was recorded.

6.3.3 Desensitization, Training and Testing procedure

Typically an acoustic bridge (a sound signal paired with primary reinforcement such as food) is used for animal training. To ensure our seals were not biased towards the fish tag due to a learned association between sound and food, the seals where not exposed to an acoustic bridge, nor was any other sound associated with the presence of food while in our care outside of the experiments reported here.

The seals were initially reluctant to approach the test boxes. Consequently each seal was given a desensitization period where they were free to access and take fish from a single box. This desensitization occurred in a separate pool adjacent to where the test trials occurred. Each seal retrieved ten fish from this single box before the experiment began.

In the learning experiment, each seal was released into the 20 box array where two boxes contained a fish. During each trial, the tagged and untagged fish where placed into two separate boxes, pseudo-randomly balanced such that throughout the course of the 20 trials every location was baited once with a tagged and once with an untagged fish. One of these, the tagged fish treatment, also contained two Vemco V9-2H coded fish tags which emitted an intermittent 69 kHz signal (source level 151 dB SPL re: 1 µPa, Fig 6.3). Each signal consisted of an eight pulse emission unique to each tag (interval between pulses ranged from 0.25 to 0.6 sec), which on average resulted in a tag signal in the pool every 13 seconds (measured across a one hour period, average inter-signal interval of 13 sec, ± 8). To monitor the tag signal, all sessions were audio recorded using a Lumbertek TS150 hydrophone and Edirol R44 recorder (sampling rate 192 kHz, 24 bit). The other box, the silent fish treatment, only contained a fish (no tags) and did not emit any sound. The seal was free to visit and revisit the boxes in any order. When the seal retrieved the tagged fish, the tags stayed in the box and predominantly continued to emit signals until the trial ended. However, in 18% of trials at the start of our experiments the reed switch was set to turn off one or both of the tags after the fish was retrieved. This was done in case a continuous tag was aversive to the animals. Once I saw no reactions to the continuing tag signals, I left tags active after fish retrieval. The
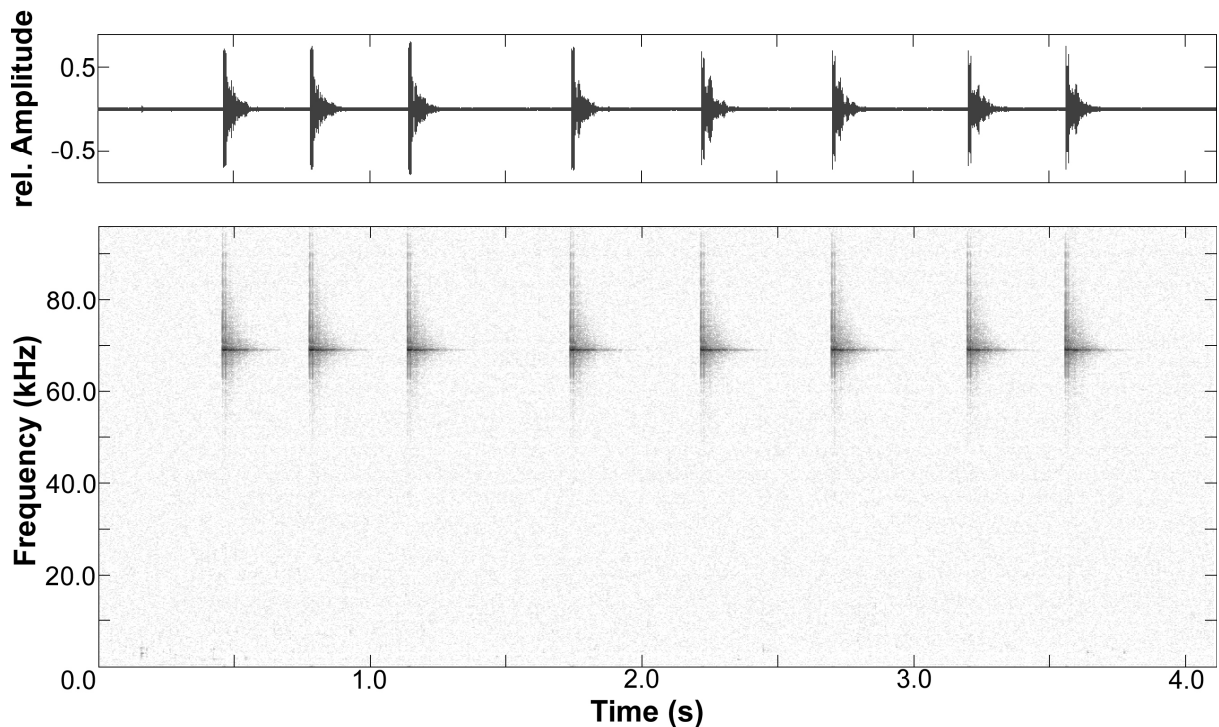
Figure 6.3. Waveform and spectrogram display of a Vemco V9 coded fish tags, emitting an intermittent 8-pulse, 69 kHz signal (source level 151 dB SPL re: 1 µPa).

trial was ended by removing the seal from the test pool, either five minutes after the seal had found both fish or after one hour if both fish were not found. Nine seals took part in 20 of these learning trials, while one animal only had 19 trials. Each seal took part in up to eight trials per day, with successive trials being a minimum of 15 minutes and maximum of 48 hours apart.

Initial results from the four animals tested in 2011 showed that the seals found both the tagged and untagged fish faster and with fewer box visits across the learning period. This suggested that the seals were at least partly using alternative cues to the acoustic cue to locate fish during the learning experiment. Hence, two additional control experiments were conducted. The first 'tag only' experiment consisted of two trials where acoustic tags were placed in one of the twenty boxes, but no fish was placed in any box. As no fish was in the pool, this eliminated any possible alternate cues. All ten seals took part in the 'tag only' trials. The second 'all fish' experiment was carried out with the six seals studied in 2012 and consisted of two test trials where additional fish pieces (inaccessible to the seal) were placed in the previously empty 18 boxes, so that each box contained either a whole accessible fish or a piece of

fish that was inaccessible at the start of the tests. Similarly to the learning experiment, only two fish were accessible to the seal (the tagged fish and the silent fish), the position of which were randomized for each trial. For inaccessible fish, the seals could still reach into the boxes with their heads, but could not reach the fish piece. Between trials the tagged and untagged fish were replaced, while the inaccessible fish were reused. During the second trial, the inaccessible fish pieces were relocated from the new accessible fish boxes to the previously used accessible fish boxes. Thus, in trial one all fish (both accessible and inaccessible) were new, while in trial two the accessible fish were new while the inaccessible fish were reused. Both test experiments occurred with a maximum delay of two days after each individual completed the learning experiment. Test trials had a maximum inter-trial interval of 20 minutes. In 2012, when seals went through both kinds of tests (the "tag only" and "all fish" control trials), the two tests were conducted with a minimum of 15 minutes and maximum of 48 hours apart.

6.3.4   Analysis

If the acoustic signal emitted by the fish tag was used as a signal for prey detection, the seals should have found the tagged fish in less time and with fewer box visits than the silent fish. During the learning period in which the association between the tag and fish was made, the time and number of box visits to finding the tagged fish should have decreased across trials. As the tags emitted sound intermittently and at random intervals, the inconsistent signal may have made the box difficult to localize. Hence, the number of repeat box visits for each box type (the box with the silent fish, with the tagged fish, and all other boxes containing neither fish nor tag) per trial was used as an additional response variable.

Data were analysed using Generalized Linear Mixed Effects Models (GLMMs) (Bolker et al., 2009; Zuur et al., 2009). GLMMs were used as the distributions of all response variables were non-normal but could be well modelled with a Poisson (repeat visits, number of boxes visited) or gamma distribution. As subjects were tested repeatedly during consecutive trails, a

crossed random effect between individual and trial number was included in the candidate models to account for repeated measures (Bolker et al., 2009). Models were fitted using the glmer function in the lme4 package for R 3.0.1 (Bates et al., 2013). The models all included at least box type (box with or without fish and/or acoustic tag) as a fixed effects factor and subject as a random effects factor. Additionally trial number and box distance from the position where the seals entered the pool were included as covariates. I also examined trial date (which would also explain age related effects as the seals were all born at approximately the same time), but this was not retained in any model.

Models were fitted to the separate data from each of the three different experiments. For the learning experiment (20 trials) I created three models; a model to predict time taken to find the fish (with a gamma error distribution and logarithmic link function), a model to predict the number of boxes visited before retrieving the fish (Poisson error distribution and logarithmic link), and a model to predict the number of repeat visits by box type (i.e. tagged fish, untagged fish and empty boxes) (also with a Poisson error distribution and log link function). Trial length was not included in the offset of this model as it did not show a correlation with the number of box visits. For both of the control experiments, I created models to predict time to fish retrieval (gamma error distribution and logarithmic link function).

In cases where a seal failed to find either the tagged or silent fish and therefore no time to fish retrieval could be measured, both observations within the trial were excluded from analysis. Additionally, for some trials in the repeat visit data set a door switch malfunctioned so that no count for the number of box visits could be obtained. 19 out of 597 trials were excluded due such door malfunctions.

A step-wise model selection procedure was carried out using a second order Akaike Information Criterion (AICc) (Zuur et al., 2009). Firstly, the 'beyond optimal model' with the interaction term (and additional covariate) was specified and different random effects combinations were tested. Secondly, the optimal combination of fixed effects was determined. Tested fixed effects included trial

number and its interaction with box type and distance from the point where the seal entered the pool to the fish boxes as a potential additional covariate (for the 'time to fish' models only). The interaction term of box type (tag presence) and trial number would indicate a learning effect (i.e. the seal finding the tagged fish faster towards the end of the 20 learning trials). In one case, a candidate model for the test experiment did not converge and had to be excluded from the selection process. In the box visit model, contrasts between the three levels of the factor box type were tested using the lsmeans function from the lsmeans package in R (Lenth, 2013).

The final model assumptions were checked using diagnostic plots of model residuals. This procedure revealed one residual which was an extreme outlier that could disproportionately influence the overall outcome of the 'time to fish' model for the learning experiment. To test the effect of this residual, the model was refitted without the outlier and these results are presented separately. Confidence intervals (CIs) were calculated using Wald statistics and both model parameter estimate and CIs are shown on the scale of the response variable.

6.4   Results

In the 20 trial learning experiment, time and number of boxes visited to finding fish decreased across training trials, showing a learning curve (Figure 6.4 and 6.5). The mixed model showed a highly significant effect of trial number with a reduction in time and number of boxes visited before retrieving the fish over consecutive trials (Table 6.1). The GLMM also showed that seals needed less time and fewer box visits to find fish in boxes which were closer to the pool entrance. The interaction term of tag presence and trial number was highly significant for number of boxes visited before fish retrieval. This indicates that seals visited fewer boxes before finding the acoustically tagged fish compared to the untagged fish in later trials. However, there was no consistent effect of box type (tagged or silent) on time needed to finding the fish (Table 6.1). This suggests that seals spontaneously used alternative sensory cues to locate fish from silent boxes and perhaps even from tagged boxes during the learning
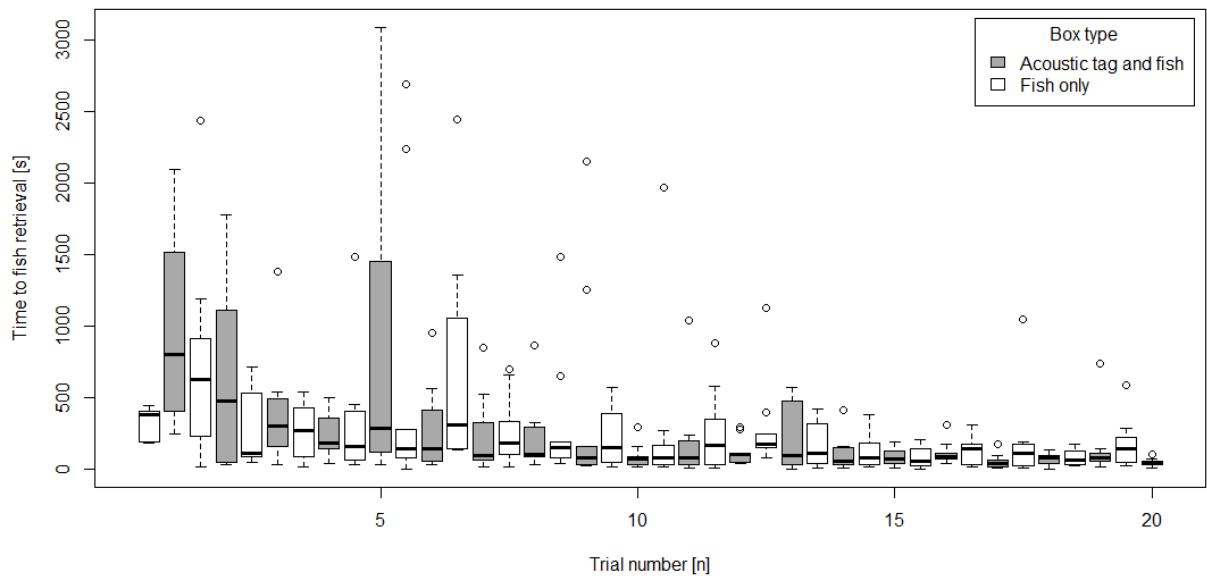
Figure 6.4. Tukey's boxplots for the time to finding fish, either with or without the fish tag, by trial throughout the learning period.
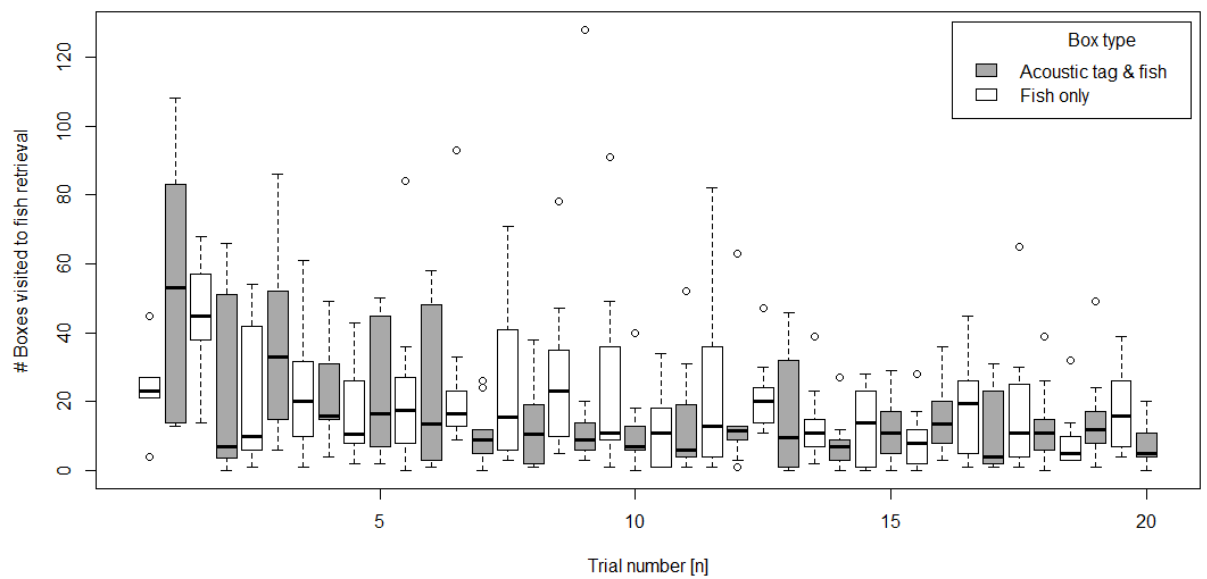
Figure 6.5. Tukey's boxplots for the number of boxes visited before finding fish, either with or without the fish tag, by trial throughout the learning period.

Table 6.1. Generalized linear mixed effects models (GLMM) for the time and number of boxes visited before finding the tagged and silent fish during the 20 learning trials (gamma distribution & log link). Model estimates for fixed effects are presented on the scale of the response variable. Significant (p<.05) variables are highlighted in bold.

	Model	Estimate e^{β}	Confidence interval		P
			2.5%	97.5%	
(Intercept)	Time	480.72	235.33	981.98	**<0.0001**
	Box visits	20.671	16.354	26.128	**<0.0001**
Acoustic tag	Time	1.245	0.738	2.101	0.412
	Box visits	0.971	0.789	1.195	0.783
Trial number	Time	0.915	0.882	0.949	**<0.0001**
	Box visits	0.957	0.940	0.975	**<0.0001**
Distance	Time	1.014	1.001	1.026	**0.0382**
	Box visits	1.020	1.017	1.023	**<0.0001**
Acoustic tag* trial number	Time	0.963	0.923	1.004	0.0792
	Box visits	0.982	0.973	0.991	**0.0001**

experiment. There was, however, some evidence for an additional learning effect related to the presence of the acoustic tag. The interaction term of box type (presence of the tag) and trial number approached significance in the standard model and became highly significant in the model when a single extreme outlier was removed (Table 6.2). This significant interaction indicates that seals needed ~5% less time to find the box which contained the tag with each consecutive trial. Additionally, the seals revisited the box with the tagged fish most frequently (Figure 6.6).

The mixed model (GLMM) for repeated box visits (table 6.3) showed that seals visited boxes which initially held the untagged fish 1.4 times more often than empty boxes. However the acoustic tag caused a 2.4 fold increase in the number of repeat visits compared to empty boxes. Seals revisited boxes with

Table 6.2. Results from the generalized linear mixed effects models (GLMM) for the time before finding the tagged and silent fish during the 20 learning trials with the outlier removed (gamma distribution & log link). Significant (p<.05) variables are highlighted in bold. Model estimates for fixed effects are presented on the scale of the response variable.

	Estimate e^{β}	Confidence interval		P
		2.5%	97.5%	
(Intercept)	445.8	231.42	858.87	**<0.0001**
Acoustic tag	1.363	0.860	2.160	0.188
Trial number	0.914	0.882	0.950	**<0.0001**
Distance	1.018	1.007	1.030	**0.001**
Acoustic tag * trial number	0.949	0.914	0.985	**0.006**
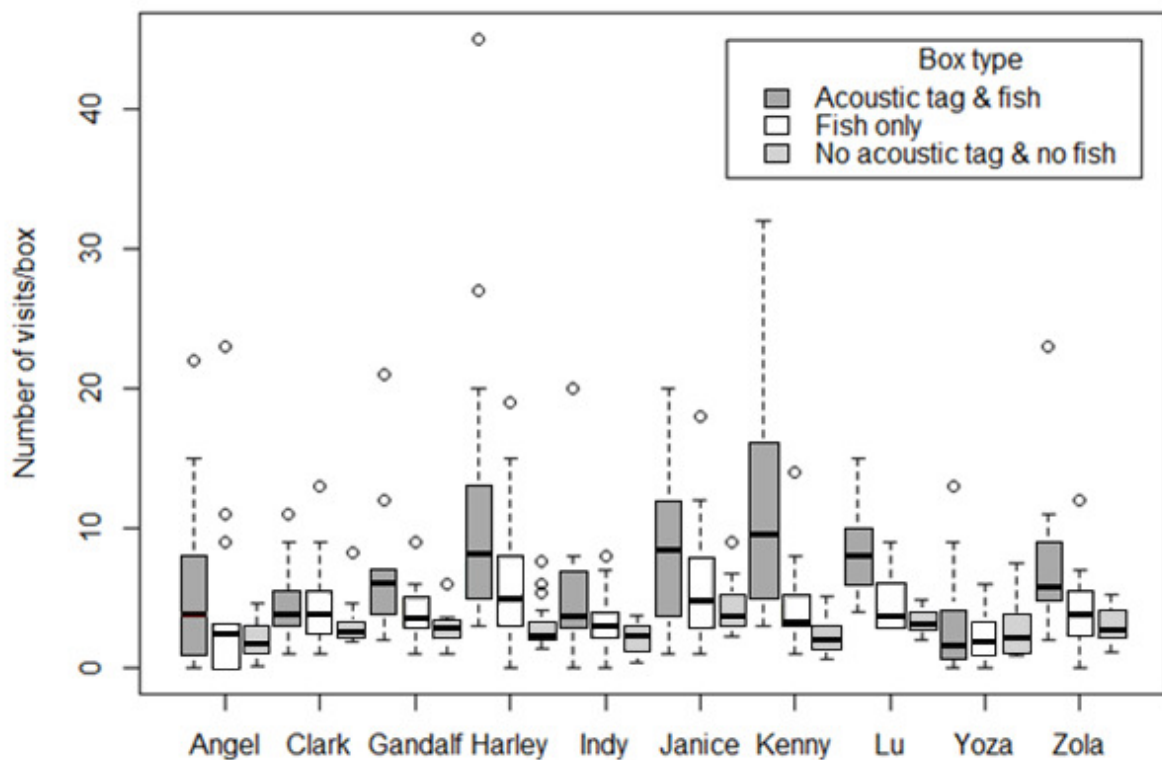
Figure 6.6. Tukey's boxplots for the number of repeat box visits by individual seals' by box type, either fish with or without the fish tag or empty box, in the learning experiment. Pairwise difference obtained from model contrasts are shown above the graph. *** = p<0.0001.

Table 6.3. Generalized linear mixed effects model (GLMM) for the number of repeat box visits during the 20 learning trials (Poisson error distribution and log link). Model estimates for fixed effects are presented as incident ratios on the scale on the response variable. Significant (p<.05) variables are highlighted in bold. 'NR' indicates the variables were not retained in the model selected.

	Model	Estimate e^{β}	Confidence interval		P
			2.5%	97.5%	
(Intercept)	All Fish	249.6	56.678	1099.232	**<0.0001**
	Tag Only	66.511	38.644	114.474	**<0.0001**
Acoustic tag	All Fish	0.011	0.003	0.044	**<0.0001**
	Tag Only	0.464	0.341	0.632	**<0.0001**
Trial number	All Fish	0.221	0.086	0.568	**0.0017**
	Tag Only	2.090	1.413	3.089	**0.0002**
Distance	All Fish	1.085	1.053	1.119	**<0.0001**
	Tag Only	NR	NR	NR	NR
Acoustic tag* trial number	All Fish	8.637	3.688	20.226	**<0.0001**
	Tag Only	NR	NR	NR	NR

the acoustic tag significantly more often than boxes which initially contained untagged fish as revealed by highly significant contrasts between the three levels of the factor box type (p<0.0001).

In the first control experiment (the 'tag only' trials), no fish was placed in any box while one box contained acoustic tags. Time to finding the acoustic tag was compared to that of a randomly selected box. The results differed markedly from the learning experiment as tag presence caused a significant reduction (~54%) in time needed to visit the box (GLMM, table 6.4, Figure 6.7), confirming the seals learned to use acoustic cues in the 20 initial trials. While tag presence reduced the time needed to find the acoustic tag across both trials, the time needed to visit a box was twice as high in the 2nd trial compared to the first (GLMM, table 6.4).

These results were supported in the second control experiment (the 'all fish' trials), in which other sensory cues were made unreliable by placing fish in all boxes, with only two of these fish being accessible to the seals. The model showed that tag presence caused a significant reduction in time needed to retrieve the fish (GLMM, table 6.4, Figure 6.8). The interaction between trial number and tag presence was also significant, showing the effect differed between the 1st and 2nd test trial. The model contrasts which show the significant differences between the pairings are presented in Figure 6.8. Seals needed less time to find the tagged fish in trial one and two when compared to finding the untagged fish in trial 1. It also becomes obvious that seals needed less time to find the silent fish in the 2nd test trial than to finding it in the 1st test trial which may indicate the ability to differentiate chemosensory cues from fresh versus reused fish. Seals also found the fish near the pool entrance faster than those further away.

Table 6.4. Generalized linear mixed effects models (GLMM) for the time to finding the tagged versus control box during the two test conditions (gamma distribution & log link). Model estimates for fixed effects are presented on the scale of the response variable. Significant (p<.05) variables are highlighted in bold. 'NR' indicates the variable was not retained in the model selected.

	Model	Estimate e^{β}	Confidence interval		P
			2.5%	97.5%	
(Intercept)	All Fish	249.6	56.678	1099.232	**<0.0001**
	Tag Only	66.511	38.644	114.474	**<0.0001**
Box type	All Fish	0.011	0.003	0.044	**<0.0001**
	Tag Only	0.464	0.341	0.632	**<0.0001**
Trial number	All Fish	0.221	0.086	0.568	**0.0017**
	Tag Only	2.090	1.413	3.089	**0.0002**
Distance	All Fish	1.085	1.053	1.119	**<0.0001**
	Tag Only	NR	NR	NR	NR
Box type* trial number	All Fish	8.637	3.688	20.226	**<0.0001**
	Tag Only	NR	NR	NR	NR
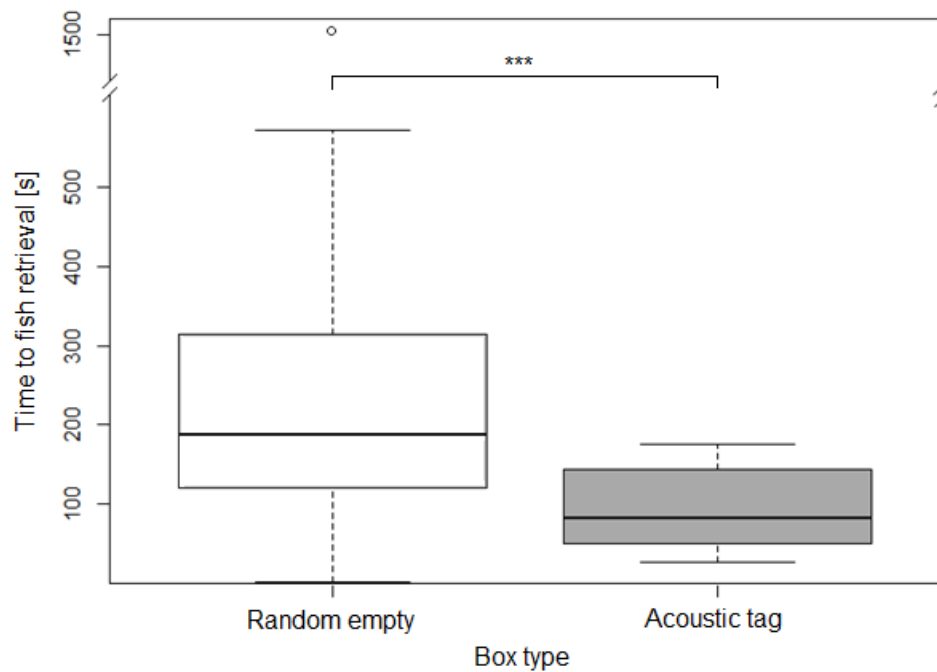
Figure 6.7. Tukey's boxplots for the time to finding the tagged box and random selected empty box for the two 'tag only' test trials. Model contrasts are shown above the graph with *** = <0.0001
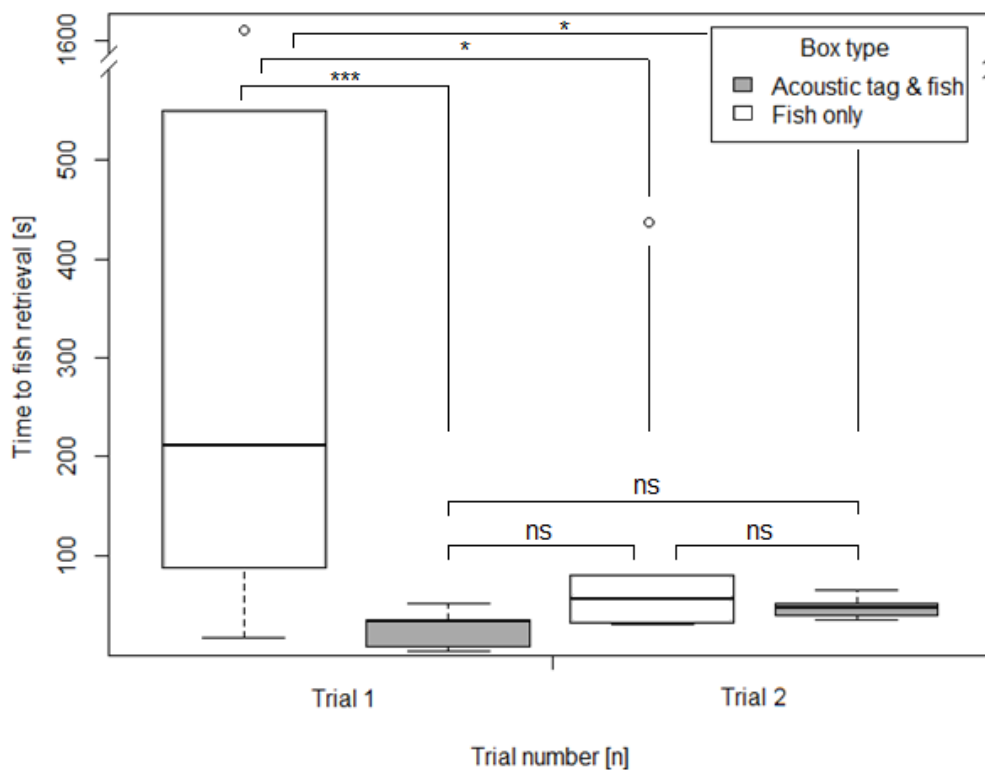
Figure 6.8. Tukey's boxplots for the time to finding the fish, either with or without the fish tag, for the two 'all fish' test trials. Model contrasts are shown above the graph with *** = <0.0001, *= <.01, and ns = not significant.

6.5 Discussion

My study documented the spontaneous use of novel environmental cues during a foraging task, showing that grey seals are capable of acoustic comprehension learning. The seals dramatically reduced the time and number of boxes visited to find fish by adapting their foraging behaviour to use environmental information within a relatively short time period. The significant interaction between trial number and box type shows that the seals found the tagged fish in fewer box visits than the untagged fish in later trials, demonstrating the learned use of the acoustic tag to locate food. While seals did not find the tagged fish much faster than the silent fish during this learning experiment, there is evidence for a weak interaction effect between trial number and presence of the acoustic tag.

These results indicate that subjects may have used alternative sensory cues, in addition to the acoustic tags, to find fish. It seems most likely that chemosensory cues were used. As seals keep their nostrils closed underwater, olfactory detection is unlikely. It is thus most probable that gustatory information was used by the seals. In the initial trials, the seals appeared to use this alternative cue as a primary source of information to locate fish during the learning experiment, but gained additional information from the tag in later trials.

The use of the acoustic tag is additionally supported by the increased number of repeat visits to tagged boxes, demonstrating that seals learnt the relevance of the acoustic cues and adjusted their foraging strategy to revisit profitable foraging spots. While the seals may have initially revisited tagged boxes as an exploration of a novel stimulus (the tag) it would be expected to decrease with experience. However, there was no difference in repeated box visits across trials (not retained in model selection). Additionally, this increase of visits to the tagged box location was stronger than to the silent fish location, showing that animals did not just return to a previously successful foraging site but instead they were influenced by the continuing acoustic signal. These results are particularly relevant when considering the potential impact of long-term sound sources, such as net pingers or ADDs. The seals were 2.4 times more likely to re-visit a tagged location than any other empty box, despite the

tagged box remaining depleted (Table 2). Once a sound is associated with prey, it can be expected that animals will repeatedly revisit that location for some time. This effect was significant in the experiment despite a number of initial trials in which the tag did not emit signals after fish retrieval (see methods).

The effect of the tag became apparent in the control experiments where alternative, presumably gustatory, cues became unreliable due to either no fish being placed in the boxes ('tag only' trials) or the presence of fish in all boxes ('all fish' trials). In both of these experiments, presence of the acoustic tag caused significant reductions in the time needed to find the tagged fish. In the 'tag only' trials the tagged box was found faster in the first trial, and it took twice as long to find in the second trial. This may be due to an extinction effect, where exposure to the tag without a fish reduced the animal's response in the second trial. Interestingly, in the 'all fish' trials, the seals also managed to reduce the time needed to find the silent fish in the 2nd test trial, suggesting a chemosensory ability to distinguish older from more recent baits since inaccessible fish was not changed between trials. Sea lions (Friedl et al., 1990) and harbour seals (Sticken & Dehnhardt, 2000) are able to detect chemosensory cues in water, but my results are the first to suggest the use of such cues during foraging.

My findings present a novel way of looking at anthropogenic signals that illustrates how animals exploit cues when they become available. It is difficult to assess the extent to which seals could rely on such acoustic and chemosensory cues present in this experiment when foraging in the wild. Live, mobile fish are likely to provide less chemosensory information than the dead fish used in our captive experiment, which could make acoustic fish tags in the wild a more dominant and reliable cue. However, the movement of live fish together with the low duty cycle of acoustic tags may make acoustic signals less efficient for detection of pelagic fish. The acoustic signal from a tag may be most beneficial to a predator when emitted from sedentary and inconspicuous prey like flat fish where hydrodynamic swim trails that can be used for prey detection (Schulte-Pelkum et al. 2007; Hanke et al., 2000; Dehnhardt et al., 2001) are absent. Detection range may impact which cues are most salient; acoustic tags may

increase prey detection from a distance by attracting experienced seals to locations with mobile tagged fish, where they then use other sensory inputs for prey capture. My results therefore illustrate the importance of considering the auditory sensitivities of all animals in the environment when designing an acoustic tagging study for a selected species. The learned association between a signal and food leading to a 'dinner bell' effect has been demonstrated in several species. Other marine animals are similarly capable of utilizing noise information and associative learning. This effect may be most pronounced in marine mammals with low auditory thresholds in high frequency bands. Detection ranges for 69 kHz tag signals in odontocetes, for example, have been predicted to exceed one km (Bowles et al., 2010).

Acoustic fish tags are being used extensively in mark-recapture studies to assess fish survival (Starr et al., 2005; Melnychuk et al., 2010; Wargo-Rub et al., 2012a & b). Research agencies worldwide invest significant resources in acoustic tagging studies to assess fish stocks and determine survival rates. As acoustic tags could make a fish more vulnerable to predation, tagging can lead to erroneous conclusions in such studies. This concern is supported by observations of decreased survivorship rates for acoustically tagged juvenile salmon compared to those with similar tags that produce no sound signal (Wargo-Rub et al., 2012a & b). Similarly, tagged predator species may experience a decrease in foraging success. Acoustic tags are becoming more widely used on sharks (Garla et al., 2006; Heupel & Simpfendorfer, 2011; Barnett et al., 2012) and could make them more detectable by prey species such as seals (Cunningham et al., 2014). Even recently published reports of acoustically tagged seals meeting at sea (Lidgard et al., 2012) could be caused by a tag attraction effect, since the tags used produced sounds similar to the fish tags used in the area. In the case of the seals, a possible solution to reduce detectability of tags would be to increase the frequency of the tags. Such tags are currently commercially available. However care should be taken as other predators with higher frequency sensitivity, such as cetaceans, could still detect such tags.

All tagging studies rely on the basic assumption that tags have no significant impact on marked individuals. However, my results suggest that acoustic tags could have profound effects on the fitness of the studied individuals in situations where they are audible to conspecifics, predators or prey. Similar tag effects have been widely investigated in the use of rings to mark birds; ring colour and symmetry alters mate selection, reproductive success (Burley, 1985; Zann, 1994; Hunt et al., 1997; Hagan & Reed, 1988; Fiske & Amundsen, 1997) and dominance interactions (Cuthill et al., 1997). Marking also increases detectability by predators; tadpoles marked with a skin staining dye are more susceptible to predation than unmarked tadpoles (Carlson & Langkilde, 2013). While most research has examined the effects of visual marking, here I showed that acoustic tags comparably aid prey detection, potentially increasing predation of tagged animals. Acknowledging such impacts of marking, both for visual and acoustic tags, is critical to research generalizing the behaviour and mortality of marked animals to natural populations.

Current research has focused on documenting the detrimental physiological effects of noise on animal fitness as it is of importance for conservation and the regulation of anthropogenic noise in natural environments. However, less consideration is being taken for how anthropogenic sounds may be beneficial to some organisms, such as to increase foraging success. Artificial noise sources are widely deployed in various anthropogenic activities or in an attempt to study or manipulate animal behaviour. Examples include the fish tags tested here, but other acoustic devices such as net pingers, echosounders, boat engines, turbines, sonar, and ADD's could similarly be exploited for beneficial information. This is particularly important to consider when such benefits may lead animals to expose themselves to higher sound levels for longer periods of time, potentially causing harm that would not have been accounted for when assuming signals would have a deterrent effect. Thus, when introducing artificial sound sources into an environment it is important to take into consideration all potential effects on local species, both detrimental and beneficial.

I demonstrated that through comprehension learning, anthropogenic signals can be spontaneously used to an animal's benefit as a signal to detect

prey. Similar results could be expected for many animal species that can perceive a systematically introduced sound signal. Future studies should consider animal learning capabilities when using such sound sources.

Chapter 7
General Discussion and Summary

7.1 Introduction

Throughout this thesis I have examined the vocal development of the grey seal. Past research concerning pinniped vocalisations has focused on documenting vocal repertoires and the behavioural context of calls. These studies have shown that pinnipeds have advanced vocal capabilities; they produced varied calls (in some cases with complexity comparable to bird song), have voluntary control over vocal plasticity allowing for innovation and variability within existing repertoires, and can produce these complex calls in specific contexts (e.g. McCulloch, 1999; Schusterman & Reichmuth, 2008; Shapiro et al., 2004). However, past work has provided little information on how pinniped calls change with development, particularly examining the effects of physical maturation and/or learned experience.

Within this thesis, I have attempted to address this knowledge gap. The vocal repertoire of grey seal pups, and how their calls develop and change over time, was documented. By examining call changes that occurred across animals of the same age, size, and sex, I gained an understanding of how seal calls change with physical development. Examining the changes that occurred within individual animals, such as the contexts in which calls are produced, changes in call structure, and how sound information is used, provided insight into the role of learning in development. By considering both maturational changes and learning, I have gained a more comprehensive understanding of how the complex vocal repertoires of pinnipeds arise.

Here, I will examine my findings in light of previous research and how these studies can be integrated to gain a better understanding of vocal development and flexibility.

7.2 The interaction between maturation and learning in call development

In pinnipeds, past research has been predominantly limited to documenting vocal repertoires in conjunction with observing behaviour. A few investigations have examined how these calls change with age, while others

demonstrated that pinnipeds were capable of contextual learning (e.g. Asselin et al., 1993; Job et al., 1995; McCulloch, 1999; Charrier et al., 2003; Shapiro et al., 2004; Sanvito et al., 2007). However, a comprehensive view of call development considering both physiology and learning has been previously lacking.

Here, I demonstrated how grey seal calls changed with age (chapter 3). In general, grey seal calls decreased in frequency, noise and variability and increased in duration and total bandwidth with age. These parameters also varied between sexes; call frequency was higher and duration shorter in females than in males. These changes can most likely be attributed to physical development, as they occurred across individuals and follow developmental patterns similar to most other mammals. These patterns would be expected given the anatomical changes that occur with maturation such as the lengthening of the vocal tract, descent of the larynx, changes in the tension of the vocal folds, and increased lung capacity (chapter 3).

In addition to examining call changes which can be attributed to physical growth, I also demonstrated that grey seals were capable of changing their sounds through learning. Grey seals can learn to produce specific sounds in specific contexts by classifying and generalizing sounds into call types (chapter 4). They can learn to manipulate the structure of their calls, allowing for imitation (chapter 5). Lastly, grey seals can also learn from the sounds they perceive and use this sound information to their benefit, such as for prey detection (chapter 6).

While maturation and learning were treated separately in this thesis, the two processes are entwined and cannot be separated when examining call development. Physiology restricts the extent to which learning can contribute to call repertoires; if an animal is incapable of physically producing a sound then learning will not allow for the sound to be used by the individual. This, however, does not mean learning may not be occurring; for example in some species female songbirds are capable of learning the same songs as males, but do not produce the songs without high levels of testosterone (Garamszegi et al., 2007).

Even if learning occurs, physiological constraints impact whether an individual will produce a learned call.

Physiological constraints in perception may similarly impact repertoire development. For example infant bat echolocation clicks increased in frequency with age (e.g. Moss, 1988; Sterbing, 2002; Monroy et al., 2011; Carter et al., 2014). This may be due to restrictions on the sound production mechanism with development, such that the infants are incapable of producing high frequency sounds. There may also be limitations to the infant's perception of calls. High frequency hearing increases with age, corresponding with the changes in sound production (Rübsamen, 1987). This process may indicate physiological limitations on learning; infant bats match the call frequency of their mothers, showing they modify their calls based on perception of their mother's vocalisations (Knörnschild, 2014). If at birth they are unable to perceive the high frequency components of their mother's calls, they would be limited to matching the lower frequencies they could perceive. With age and development of the auditory system, they then can utilize production learning to produce adult echolocation. Self-feedback may also explain these results if infant bats can only produce and modify their frequencies based on what they hear of their own calls. It is possible that a combination of factors, from the infant's production and perception to self-feedback mechanisms, contribute to call development. In this way, perceptual limitations mediate what sounds an individual can hear, and thus limit what can be learned.

Learning may also mask changes occurring during physical development. Through learning, I have shown that animals may produce specific call types with set call parameters. Although individuals may be capable of producing other sounds with a greater range of call parameters, individuals may only produce a selection of these calls based on learned experience. For example, I observed that captive juvenile grey seals produced fewer calls than they did as pups. This could be attributed to a developmental change, where animals produce fewer calls with age, or as an artefact of being in captivity.

Alternatively, I could have artificially trained a low sound production rate. Some of the seals in our study were rewarded for producing sound when cued,

and for refraining from producing sound when that cue was not present. As cues were not provided outside of training sessions, this could have resulted in an abnormally low call rate. Interestingly this did not seem to be the case; outside of training sessions, the seals who took part in training produced more calls than those who were not trained. In this case, training may have encouraged increased call rate. Similarly, training may have inadvertently rewarded the seals for producing calls with certain acoustic parameters during sessions, such as certain durations or frequencies. Through operant conditioning, this may have encouraged the seals to have produced those parameters more frequently even outside of training sessions. Learning may have thus altered the individual's behaviours, and would misrepresent average developmental trends.

Given this complex interaction between physiology and learning in vocal development, these processes cannot be separated when examining normal, healthy individuals. Changes in both physiology and learning should ideally be considered concurrently. Sounds are complex stimuli and when perceiving calls of other individuals, receivers may gain separate information from call parameters driven by physical and learned changes. For example, hyrax calls can simultaneously encode information concerning age, weight, size, physical condition, hierarchical status and hormonal states (Koren & Geffen, 2009). Interestingly, I found that the frequency parameters of the seal calls could provide information in both physiologically constrained parameters as well as learned ones. During development, the frequency parameters of seal calls decreased with age, and were lower in males than females. This suggests that call frequency could be used as an honest indicator of the caller's age, size and/or sex (chapter 3). When examining usage learning, I found that frequency parameters could also be used to distinguish individual identity (chapter 4). This information regarding individual, age, size and sex is most likely physically constrained. However, when testing usage learning I also observed that frequency parameters could be used to distinguish between call types (chapter 4), and when examining vocal production learning I found that individuals could learn to alter their own calls to match the peak frequencies of models (chapter

5). In these cases, animals learned to alter different frequency parameters. While some frequency parameters were physically constrained, others could be altered with learning, and thus information about both the caller and learned content could be simultaneously encoded in calls.

In this way, learning to alter frequency information does not necessarily preclude simultaneous honest signalling. In production learning, the seals did not match fundamental frequency of signals. Instead they matched the peak frequencies/formants of the call. This might suggest that callers can provide honest biological information such as age, size, sex and identity in the more certain aspects of the call, such as the fundamental frequency, in addition to encoding other information within the peak frequencies of the call. The seals may also be limited by physiology in the range of sounds they can produce. So while they may learn to alter frequency parameters, this can only occur within a set range depending on physical restrictions. Alternatively, in some cases individuals could adjust call frequencies with learning to provide dishonest signals. This might be desirable in cases such as mate selection; males may be selected by females for size, and males who learn to make themselves sound 'bigger' may be more likely to be chosen by females.

This balance between fixed and flexible aspects of sound is particularly interesting when considering that in many cases animals may favour more stable call structures in order for information between sender and receiver to be correctly interpreted. In such cases, vocal learning may serve to help individuals produce calls which more closely match an auditory template and to thus function for communication. As previously discussed, calls are multi-dimensional stimuli, and perhaps while certain call features are more constrained, others are flexible and allow for learned changes. Future research more closely examining where call properties remain fixed, where there is flexibility, and how these parameters interact to provide information is needed.

7.3   Vocal Learning and Language

Sounds are complex stimuli in that they can provide a wide range of information within a single signal. This is apparent in human language; when I

speak, I provide information about my identity, such as sex and age, in addition to the learned meaning of the words. This may also occur to an extent in animals; for example, vervet monkeys (*Cercopithecus aethiops*) produce predator alarm calls which may have learned contextual meaning (Seyfarth et al., 1980a) but can also indicate individual identity, age, and sex (Cheney & Seyfarth, 1980 & 1988; Hauser, 1989).

Despite some similarities, there is a distinctive gap between the advanced spoken language of humans and the vocal abilities of other animal species. These differences have led to language being defined as being uniquely human (Chomsky, 1957; Hackett, 1960; Macphail, 1982; Pearce, 1987). Hackett (1960) described the qualities that make up human language, referred to as 'design features', several of which appear to be lacking in animal communication systems. For example, many animal vocalisations lack semanticity (a set meaning for a specific sound signal), discreteness (sounds consist of discrete units, and by changing the order of these units the meaning is also changed), and productivity (sounds have infinite potential to indicate different meanings, through the development of new call types and by changing how the units are organized).

While all of these abilities are present in human communication systems, in combination they have not been demonstrated in any non-human mammal. However, this may be attributed to a lack of relevant studies; for example cetaceans show semanticity and discrete structures in their vocal communication (e.g. Harley et al., 2005; Suzuki et al., 1999), however no study has documented vocal innovation in a cetacean. It seems very unlikely that cetaceans are incapable of innovation, particularly bottlenose dolphins which have been subjectively reported to produce varied calls when trained (e.g. Richards et al., 1984; Harley et al., 2005). It is also very likely that changes in humpback whale song can be attributed to innovation (e.g. Noad et al., 2000). Further investigation is warranted in such species with advanced vocal capabilities.

Although these skills are lacking in the majority of animal vocal communication systems, they may be apparent outside of the vocal domain. For

example, nonhuman primates are capable of using sign language and keyboard based communication systems (e.g. Gardner & Gardner, 1969; Premack & Premack, 1972; Patterson, 1978; Terrace et al., 1979; Williams et al., 1997). However, the repertoire of these non-verbal cues appears to be relatively small, limited to the scale of hundreds of words, in comparison to the arguably infinite human language.

Within this thesis, I demonstrated that seals were capable of controlling when they produced specific calls, changing call structure to match models, and learning to use information from perceived sounds. While arguably advanced, these abilities only demonstrate some of the fundamentals necessary for the development of language, and do not suggest that seals are capable of communication comparable to human speech, despite the observation of the harbour seal, Hoover, who copied language sounds (Ralls et al., 1985). We are only beginning to grasp the complexities of animal communication systems, and by doing so understanding the evolution of human language in relation to the vocal abilities of other species. This thesis illustrates some of the advanced capabilities seals may possess, but does not assess the extent of their abilities. Further research into the capabilities and limitations of pinniped communication would be valuable to gain a better understanding of how flexible their vocal system is in comparison to other species, including humans. For example, would a seal be capable of learning the correct usage of human speech sounds, such as has been demonstrated in parrots (Pepperberg, 1999)? Although this thesis touched on some of these questions, particular areas of interest for future research examining vocal abilities in pinnipeds include:

- What are the physical limits for sound production and perception?
- What are the natural units of sound used in call production? Can arbitrary units be learned, or are these units fixed from birth?
- How are sound units naturally perceived, and how are they classified as being the same or different? Can arbitrary classifications of these units be learned?

- Are there syntactical rules to how sound units are naturally used together? Can new rules be learned?
- Are sound units used referentially? Do syntactical changes alter referential meaning?
- How flexible are sound production systems? Is there potential for novel units to be developed or rearranged syntactically?

7.4 Comparative Perspectives on the Social Functions of Vocal Production Learning

I chose to examine seals because past studies have shown they have diverse call repertoires, have voluntary control over vocal plasticity, are capable of contextual vocal learning (e.g. Shapiro et al., 2004; Schusterman, 2008; Sanvito et al., 2007) and anecdotal evidence suggests they are capable of vocal production learning (Ralls et al., 1985). However, pinnipeds are an interesting model for vocal learning given their fluid social structure. Many pinniped species are thought to be relatively isolated for the majority of their life while travelling at sea. However, individuals regularly come together in dense aggregations when hauled out on land, particularly during breeding and pupping seasons. Vocalisations are thought to play a role in social interactions during these periods, allowing for individual identification, mother-pup reunions, mediating agonistic and/or territorial interactions, and facilitating mate attraction.

Other species with advanced vocal learning capabilities also have complex social structures. All of the other accepted mammalian species capable of vocal production learning (cetaceans, elephants and bats) have fission-fusion societal structures with individuals moving dynamically to form different groups over time (e.g. Lusseau et al., 2003; Archie et al., 2006; Kerth et al., 2011). Across species, increased social variability appears to positively correlate with vocal repertoire size (Blumstein & Armitage, 1997; McComb & Semple, 2005; Freeberg, 2006; Le Roux et al., 2009). Perhaps advanced vocal learning abilities would be expected of species with fluid social structures given that the most likely functions of vocal learning appear to be social. Without rigid

social structures, learned vocal signals may be more important for mediating social interactions.

This is particularly interesting when considering pinnipeds; individuals presumably spend the majority of time isolated at sea but haul out in dense aggregations for short periods, sometimes at the same sites in which case they would presumably be near known individuals (Pomeroy et al., 1994; Pomeroy et al., 2000). Other vocal learning species have much longer social relationships, lasting for years at a time (e.g. Wells, 1991; Reynolds et al., 2000). The duration of such associations may influence how production learning is used by these species. Cetaceans are arguably amongst the best non-human mammalian vocal learners, and they have long-term social relationships which change within a group over time (e.g. Wells, 1991; Reynolds et al., 2000). Although pinnipeds may haul out with the same individuals regularly (Pomeroy et al., 1994; Pomeroy et al., 2000) in comparison they do not have such a complex social structure. While pinnipeds are capable of production learning, their abilities do not appear to be as striking as those of cetaceans. Perhaps this can be attributed to having a simpler social structure, with shorter-term interactions, which does not necessitate advanced production learning abilities. Future research examining vocal learning may benefit from focussing on the vocal capabilities of species with dynamic social structures, particularly considering the strength of social associations.

However, not all species with flexible social structures are also capable of vocal learning. Many non-human primates also have fission-fusion social structures (Symington, 1990), but despite extensive study appear to have rigid vocal repertoires. Perhaps while vocal production learning is more likely to occur in species with flexible social structures, not all of these species will necessarily be capable of vocal learning. Alternatively, perhaps the vocal abilities of primate species have been underestimated and further examination may be warranted (e.g. Crockford et al., 2004; Candiotti et al., 2012; Watson et al., 2015).

Pinnipeds are also interesting given that several species produce song during mating season (e.g. Green & Burton, 1988b; Morrice et al., 1994). It is

assumed that males produce these songs (e.g. Green & Burton, 1988b; Morrice et al., 1994), however, vocal production learning in mammals does not appear to be limited by sex (e.g. Jones & Ransome, 1993; Miksis et al., 2002; Foote et al., 2006). Here, I demonstrated vocal production learning for both male and female grey seals. Perhaps while some pinniped species use vocal production learning to facilitate male song production for inter- or intra- sexual selection, the ability may have been retained allowing for production learning in females and other pinniped species which may not produce song.

Here I demonstrated that grey seals are capable of vocal production learning through imitation of novel sound signals. However, are grey seals unique in their vocal learning abilities? The majority of pinniped vocal research has focused on phocid species, which are assumed to be vocal learners. Fewer studies have investigated otariid species, and they are currently thought to be incapable of vocal production learning. Given that so few studies have systematically examined vocal learning in pinnipeds, it is unclear as to how widespread their capabilities are. Is the capacity for vocal production learning unique to a few species such as the grey and harbour seal, or is it common among all phocid seals? Alternatively, are all pinniped species capable to some degree of production learning?

Similarly, are other mammalian species capable of vocal production learning? Vocal production learning has only been conclusively demonstrated in humans, cetaceans, elephants, phocid seals and bats. Is this due to true differences in the capabilities between mammalian species, or can it be attributed to the way we currently measure vocal learning? Past reviews have argued that production learning may be common to a much wider taxonomic distribution than previously assumed (Janik & Slater, 1997 & 2000; Tyack, 2008). Vocal production learning is difficult to demonstrate as other factors such as genetics, maturation, improvisation, motivation and environmental changes may also explain observed repertoire changes. Convincing evidence for production learning relies upon demonstrating the animals learn a novel signal that did not previously exist in their repertoire. In many cases it is difficult to

prove that a sound is truly novel unless it is far outside the typical species repertoire, such as human language.

For the most part, attempts at teaching animals such artificial signals have been unsuccessful. This negative result might be expected, given that such artificial sounds are far outside of normal call repertoires. It is possible that for animals to learn, sound models may have to be perceived as biologically relevant. If vocal learning primarily functions for social purposes, it might be expected that only species-specific models would be copied. However, this is a dilemma given that evidence of learning from species-specific call models may not be convincing, as the calls could have arguably pre-existed in the animal's repertoire and were simply learned to be produced in a novel context as opposed to learning a new call.

In chapter 2, I reviewed current studies documenting vocal production learning. While accepted vocal learners have shown arguably advanced mimicry of biologically abnormal sound signals, it is equally important to consider subtler cases of imitation, convergence and divergence. Similarly, in addition to examining changes parameters controlled by the source and filter, parameters controlled by the respiratory system should also be evaluated. Given more flexible definitions a wider taxonomic range of species are found capable of some degree of production learning, which may provide insight into the evolution of more advanced capabilities.

In this thesis I implemented a new method of testing vocal production learning using biologically relevant sounds in a trained task. By using a positive reinforcement training paradigm, I showed that the animal clearly had learned to match playback stimuli in a consistent motivational state. By recording the seals calls and digitally modifying them to extend beyond the normal repertoire of the animals I was able to present them with novel stimuli for matching. Although not done within the scope of this thesis, extending the playback stimuli outside that of normal repertoires also enables measurement of what sounds the animals could reproduce and find thresholds for parameter limits. However, the calls were arguably still biologically relevant as the manipulated call parameters still fell within the range observed in other grey seal's repertoires. While it is still

possible that these calls pre-existed in the individual seal's repertoire, I additionally trained them to copy novel combinations of calls rather than matching single sounds. As such "program-level" imitation (Byrne, 2002) allowed for potentially pre-existing calls from the seals repertoire to be combined into novel 'songs', this leaves a question as to whether the seal demonstrated usage or production learning. The animal may have simply learned a new order to produce calls, in which case usage learning occurred, or the animal may have perceived an entire 'song' as a novel signal, in which case production learning occurred. Further studies investigating the seals perception of call units are necessary to determine if usage or production learning was more likely to have occurred.

This method also allowed for multiple acoustical parameters to be tested. Vocal learning is often viewed as a black or white capability. The ability to copy frequency is arguably advanced as it requires control over the muscles and resonance cavities within the vocal tract. Although often not considered, other species are capable of controlling a simpler aspect of the vocal tract: the respiratory system (Janik & Slater, 1997). Equally compelling evidence for vocal production learning can be demonstrated by copying call parameters controlled by respiration, such as duration, amplitude, number of calls produced, or call rate. By considering these parameters, more mammalian species would be considered vocal learners. However, there is considerable difference in the capability to imitate complex frequency changes, such as shown in the currently accepted vocal learners, and the ability to match parameters controlled by the respiratory system. To better reflect this, such categorical systems of labelling species as vocal learners or non-learners should be avoided. Instead, we should discuss vocal production learning abilities on a spectrum of varying control over call complexity. Further research examining vocal production learning of multiple parameters would be valuable, especially of parameters controlled by the respiratory system, when comparing abilities across species.

7.5   Does vocal learning occur in wild populations?

The majority of this thesis has examined trained, captive animal behaviour. Working with the animals in a captive environment was advantageous as it allowed for specific individuals to be tested over time, motivated with food rewards consistently, and with relative control over the environment. However, it should be kept in mind that when testing animals in captivity their abilities may not be representative of wild populations. Although I have shown that captive seals are capable of learning advanced vocal skills these abilities may not be utilized by wild individuals. In natural populations, animals may not need such vocal complexity.

Further studies examining the vocal development of wild populations would be valuable. When examining call development, I only examined one grey seal population on the Isle of May. Examining other grey seal sites to see if developmental trends are consistent across populations would be valuable. Additionally, while I had large recording samples taken from certain age groups, I obtained few samples of juveniles during their first year. Only a few animals were brought into captivity during this time period, and they produced relatively few calls. This may be an artefact of captivity, or call rate may normally decrease during this developmental time window. Additionally, no recordings were made of animals between 1 year and sexual maturity. These animals typically remain at the edge of the seal colony, where I was unable to access them without disturbing animals. Future research examining call development from weaning to sexual maturity of wild individuals would be beneficial.

In this thesis, learning was only examined in captive animals. Such trained matching of call types is not likely to be representative of natural behavioural states; while in a training context call matching is food reward motivated, in other species call matching appears to serve a range of functions from territorial aggression to affiliative social interactions. Call type copying has been examined in wild populations of elephant seals (*Mirounga leonina*), where juvenile males appear to produce the call types of dominant adult males preferentially (Sanvito et al., 2007). In this case, matching of dominant male seal call types is most likely used to further social status, mediate interactions

with other individuals, or for mate attraction. However, very little is known about the role of contextual learning in wild populations and further investigation of natural call repertoires, call type categorization and their use would be beneficial.

The training paradigm used to test vocal production learning in captivity is artificial and it is unclear if imitation or mimicry occurs in wild populations. Some anecdotal evidence from the research conducted on the Isle of May suggests that wild seals may naturally use vocal production learning. Firstly, during data collection for chapter 3, one adult female grey seal was observed producing a unique in-air call: a noisy pulsed sound that did not fit into previous descriptions of the grey seal's repertoires. For our analysis, this sound was classified as a 'noisy' call (examples of this call are available in Appendix XIX). Experts who have extensive experience working with grey seals and examining their vocalisations had not previously seen this call type (personal communication Simon Moss, Susanne McCulloch, Thomas Götz). The sound is somewhat similar to underwater calls types (such as the 'trrot'), which may indicate that the novel sound is a pre-existing call type being inappropriately used in air. Alternatively, subjective observations suggest that the seal may have produced this sound as a by-product of a physical abnormality in the jaw or nasal cavity while resting, perhaps comparable to human snoring.

Interestingly, a pup raised near this adult female (not her own pup) began to produce a similar sound after weaning. The call appears similar in structure to the adult's call, though the noisy pulses were higher in peak frequency. This is consistent with my results from examining call development; weaned pups had higher frequency calls than adults. The following year, the same adult female was observed producing the abnormal call type, and three new pups from the same area were recorded producing the same call upon weaning. All of the animals produced this sound in the same behavioural context (i.e. while sitting stationary, making very smalls movements with the eyes open). While unclear, it is possible that a presumably new call type may have originated in an adult female, and through learning has been socially transmitted to other individuals. Future research may benefit from examining the prevalence of this

call type within the Isle of May colony across breeding seasons, particularly when the weaned pups exposed to this adult female reach sexual maturity.

Other evidence suggesting that wild grey seal pups are capable of vocal learning is available from another unpublished study conducted on the Isle of May. This study investigated if wild seals were similarly capable of imitating digitally altered call sequences as the captive animals discussed in this thesis. Two to three 'note' call sequences with digitally altered frequency parameters (similar to chapter 5) were played back to wild nursing pups daily. There was no explicit training; the pups were simply exposed to the sound 50 times per day throughout the nursing period, starting at the age of four days old. Initial examination of the data indicates that seal pups may have adopted these novel call sequences into their repertoire and that with age these copies improved to become more similar to the example. However, these data have only been examined cursorily, and further analysis is needed to examine if production learning occurred in this case. Future research could further examine the conditions necessary for production learning in wild seals, such as whether there are critical developmental periods for learning or whether particular sound parameters are more likely to be learned.

The ability for seals to learn information from sound signals, acoustic comprehension learning, was also only examined in an artificial captive environment. A simulated foraging environment was created, where seals could use acoustic cues to find where food was hidden in possible foraging locations. While the captive seals learned to use an acoustic fish tag to find hidden food relatively quickly, it is unclear whether animals would be similarly capable of doing so with wild tagged fish. Learning would be expected to be affected by the odds of an animal encountering tagged prey, which would vary depending on factors such as the number of tagged versus untagged individuals, size of the ecosystem, and how often the tag emits sound. These effects could be estimated in future studies by modelling the probable encounters in wild populations, and would allow for further assessment of the impact of acoustic tags. Further research investigating comprehension learning in wild populations would be valuable and might reveal a wide use of acoustic cues by seals.

References

Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice*, *13* (3), 424-446. doi: 10.1016/S0892-1997(99)80048-4

Adret, P. (1992). Vocal learning induced with operant techniques: an overview. *Netherlands Journal of Zoology*, *43* (1), 125-142. doi: 10.1163/156854293X00250

Adret, P. (1993). Operant conditioning, song learning and imprinting to taped song in the zebra finch. *Animal Behaviour*, *46* (1), 149-159. doi: 10.1006/anbe.1993.1170

Aitken, P. G., & Wilson, W. A. (1979). Discriminative vocal conditioning in rhesus monkeys: evidence for volitional control?. *Brain and Language*, *8* (2), 227-240. doi: 10.1016/0093-934X(79)90051-8

Aitkin, L., Nelson, J., & Shepherd, R. (1996). Development of hearing and vocalisation in a marsupial, the northern quoll, *Dasyurus hallucatus*. *Journal of Experimental Zoology*, *276* (6), 394-402. doi: 10.1002/(SICI)1097-010X(19961215)276:6 <394::AID-JEZ3>3.0.CO;2-I

Allin, J. T., & Banks, E. M. (1971). Effects of temperature on ultrasound production by infant albino rats. *Developmental Psychobiology*, *4* (2), 149-156. doi: 10.1002/dev.420040206

Almonte, C. (2014). Classification of captive North American river otters (*Lontra canadensis*) vocal repertoires: Individual variations, and age class comparisons. *Animal Behaviour and Cognition*, *1* (4), 502-517. doi: 10.12966/abc/11/07/2014

Altmann, J. (1987). Life span aspects of reproduction and parental care in anthropoid primates. In: *Parenting Across the Life Span: Biosocial Dimensions*, (pp. 15-29). Transactions Publishers, US.

Alves, A., Antunes, R., Bird, A., Tyack, P.L., & Miller, P.J.O. (2014). Vocal matching of naval sonar signals by long-finned pilot whales (*Globicephala melas*). *Marine Mammal Science*, *30* (3), 1248-1257. doi: 10.1111/mms.12099

Anderson, S.S., & Fedak, M.A. (1987). Grey seal, *Halichoerus grypus*, energetics: females invest more in male offspring. *Journal of Zoology*, *211* (4), 666-679. doi: 10.1111/j.1469-7998.1987.tb04478.x.

Andrew, R. J. (1998). Cyclicity in speech derived from call repetition rather than from intrinsic cyclicity of ingestion. *Behavioral and Brain Sciences*, *21* (4)**,** 513–514. doi: 10.1017/S0140525X98241264

Arch-Tirado, E., McCowan, B., Saltijeral-Oaxaca, J., de Coronado, I. Z., & Licona-Bonilla, J. (2000). Development of isolation-induced vocal behaviour in normal-hearing and deafened guinea pig infants. *Journal of Speech, Language, and Hearing Research*, *43* (2), 432-440. doi: 10.1044/jslhr.4302.432

Archie, E. A., Moss, C. J., & Alberts, S. C. (2006). The ties that bind: genetic relatedness predicts the fission and fusion of social groups in wild African elephants. *Proceedings of the Royal Society B: Biological Sciences*, *273* (1586), 513-522. doi:10.1098/rspb.2005.3361.

Arriaga, G., Zhou, E. P., & Jarvis, E. D. (2012). Of mice, birds, and men: the mouse ultrasonic song system has some features similar to humans and song-learning birds. *PLoS One, 7* (10), e46610. doi:10.1371/journal.pone.0046610.

Asselin, S., Hammill, M. O., & Barrette, C. (1993). Underwater vocalisations of ice breeding grey seals. *Canadian Journal of Zoology*, *71* (11), 2211-2219. doi: 10.1139/z93-310

Atkinson, S. (1997). Reproductive biology of seals. *Reviews of Reproduction*, *2* (3), 175-194. doi: 10.1530/ror.0.0020175

Au, W.W. (1980). Echolocation signals of the Atlantic bottlenose dolphin (*Tursiops truncatus*) in open waters. In: *Animal Sonar Systems* (pp. 251-282). Springer, US.

Audacity Team (2012). Audacity, Version 1.3. http://audacity.sourceforge.net/

Balaban, E. (1994). Sex differences in sounds and their causes. In: *The Differences Between Sexes* (pp. 243-272). Cambridge University Press, UK.

Ballintijn, M. R., & ten Cate, C. (1997). Vocal development and its differentiation in a non-songbird: the collared dove (*Streptopelia decaocto*). *Behaviour*, *134* (7-8), 595-621. doi: 10.1163/156853997X00548

Baptista, L. F., & Petrinovich, L. (1986). Song development in the white-crowned sparrow: social factors and sex differences. *Animal Behaviour*, *34* (5), 1359-1371. doi: 10.1016/S0003-3472(86)80207-X

Baptista, L. F., & Gaunt, S. L. (1997). Social interaction and vocal development in birds. In *Social Influences on Vocal Development* (pp. 23-40). Cambridge University Press, UK.

Barber, J. R., Crooks, K. R., & Fristrup, K. M. (2010). The costs of chronic noise exposure for terrestrial organisms. *Trends in Ecology & Evolution*, *25* (3), 180-189. doi: 10.1016/j.tree.2009.08.002

Barnett, A., Abrantes, K. G., Seymour, J., & Fitzpatrick, R. (2012). Residency and spatial use by reef sharks of an isolated seamount and its implications for conservation. *PLoS One*, *7* (5), e36574. doi: 10.1371/journal.pone.0036574

Barton, K. (2013). MuMIn: Multi-model inference. R package version 1.9.13. http://CRAN.R-project.org/package=MuMIn

Bartsh, S. S., Johnston, S. D., & Siniff, D. B. (1992). Territorial behavior and breeding frequency of male Weddell seals (*Leptonychotes weddelli*) in relation to age, size, and concentrations of serum testosterone and cortisol. *Canadian Journal of Zoology*, *70* (4), 680-692. doi: 10.1139/z92-102

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.0-4. http://CRAN.R-project.org/package=lme4

Bates, H. W. (1862). Contributions to an insect fauna of the Amazon valley (Lepidoptera: Heliconidae). *Biological Journal of the Linnean Society*, *16* (1), 41-54. doi: 10.1111/j.1095-8312.1981.tb01842.x

Baylis, J. R. (1982). Avian vocal mimicry: its function and evolution. *Acoustic Communication in Birds*, *2*, 51-83.

Beecher, M. D., Campbell, S. E., & Burt, J. M. (1994). Song perception in the song sparrow: birds classify by song type but not by singer. *Animal Behaviour*, *47* (6), 1343-1351. doi: 10.1006/anbe.1994.1182

Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution*, *20* (3), 143-149. doi: 10.1016/j.tree.2005.01.004

Bender, R., & Lange, S. (2001). Adjusting for multiple testing—when and how?. *Journal of Clinical Epidemiology*, *54* (4), 343-349. doi: 10.1016/S0895-4356(00)00314-0

Berta, A., & Churchill, M. (2012). Pinniped taxonomy: review of currently recognized species and subspecies, and evidence used for their description. *Mammal Review*, *42* (3), 207-234. doi: 10.111/j.1365-2907.2011.00193.x

Bloomfield, L. L., Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (2003). Open-ended categorization of chick-a-dee calls by black-capped chickadees (*Poecile atricapilla*). *Journal of Comparative Psychology*, *117* (3), 290 - 301. doi: 10.1037/0735-7036.117.3.290

Blumstein, D.T., & Armitage, K.B. (1997). Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *American Naturalist, 150,* 179-200.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences*, *17* (1193), 97-110.

Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program]. Version 5.3.51, http://www.praat.org/.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24* (3), 127-135. doi: 10.1016/j.tree.2008.10.008

Boness, D. J., & Bowen, W. D. (1996). The evolution of maternal care in pinnipeds. *Bioscience*, *49* (9) 645-654.

Bordino, P., Kraus, S., Albareda, D., Fazio, A., Palmerio, A., Mendez, M., & Botta, S. (2002). Reducing incidental mortality of Franciscana dolphin *Pontoporia blainvillei* with acoustic warning devices attached to fishing nets. *Marine Mammal Science*, *18* (4), 833-842. doi: 10.1111/j.1748-7692.2002.tb01076.x

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2012). Age-and sex-specific patterns of vocal behavior in De Brazza's monkeys (*Cercopithecus neglectus*). *American Journal of Primatology*, *74* (1), 12-28. doi: 10.1002/ajp.21002

Boughman, J. W. (1998). Vocal learning by greater spear–nosed bats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265* (1392), 227-233. doi: 10.1098/rspb.1998.0286

Boughman, J. W., & Moss, C. F. (2003). Social sounds: vocal learning and development of mammal and bird sounds. In: *Acoustic Communication* (pp. 138-224). Springer, US.

Bowen, W.D. (1991). Behavioural ecology of pinniped neonates. In: *The Behaviour of Pinnipeds* (pp. 66-127). Chapman and Hill, UK.

Bowen, W. D., Boness, D. J., & Oftedal, O. T. (1987). Mass transfer from mother to pup and subsequent mass loss by the weaned pup in the hooded seal, *Cystophora cristata*. *Canadian Journal of Zoology*, *65* (1), 1-8. doi: 10.1139/z87-001

Bowen, W. D., Stobo, W. T., & Smith, S. J. (1992). Mass changes of grey seal, *Halichoerus grypus,* pups on Sable Island: differential maternal investment reconsidered. *Journal of Zoology*, *227* (4), 607-622. doi: 10.1111/j.1469-7998.1992.tb04418.x

Bowles, A. E., Denes, S. L., & Shane, M. A. (2010). Acoustic characteristics of ultrasonic coded transmitters for fishery applications: Could marine mammals hear them? *The Journal of the Acoustical Society of America*, *128* (5), 3223-3231. doi: 10.1121/1.3493438

Briefer, E.F. (2012). Vocal expression of emotions in mammals: Mechanisms of production and evidence. *Journal of Zoology*, *288* (1), 1-20. doi: 10.1111/j.1469-7998.2012.00920.x

Briefer, E., & McElligott, A. G. (2011). Indicators of age, body size and sex in goat kid calls revealed using the source–filter theory. *Applied Animal Behaviour Science*, *133* (3), 175-185. doi: 10.1016/j.applanim.2011.05.012

Briefer, E.F., & McElligott, A.G. (2012). Social effects on vocal ontogeny in an ungulate, the goat, *Capra hircus*. *Animal Behaviour*, *83* (4), 991-1000. doi: 10.1016/j.anbehavr.2012.01.020

Briefer, E. F., Tettamanti, F., & McElligott, A. G. (2015). Emotions in goats: mapping physiological, behavioural and vocal profiles. *Animal Behaviour*, *99*, 131-143. doi: 10.1016/j.anbehav.2014.11.002

Brooks, D.I., & Cook, R.G. (2010). Chord discrimination by pigeons. *Music Perception: An Interdisciplinary Journal*, *27*, 183-196. doi: 10.1525/mp.2010.27.3.183

Brown, P. E., Brown, T. W., & Grinnell, A. D. (1983). Echolocation, development, and vocal communication in the lesser bulldog bat, *Noctilio albiventris*. *Behavioral Ecology and Sociobiology*, *13* (4), 287-298. doi: 10.1007/BF00299676

Brown, E.D., & Farabaugh, S.M. (1997). What birds with complex social relationships can tell us about vocal learning: vocal sharing in avian groups. In: *Social Influences on Vocal Development* (pp. 98-127) Cambridge University Press, UK.

Buchler, E. R. (1980). The development of flight, foraging, and echolocation in the little brown bat (*Myotis lucifugus*). *Behavioral Ecology and Sociobiology, 6* (3), 211-218. doi: 10.1007/BF00569202

Bürgin, R. & Ritschard, G. (2012). Categorical parallel coordinate plot. In *LaCOSA Lausanne Conference On Sequence Analysis*, University of Lausanne, June 6th-8th 2012 (Poster), http://www3.unil.ch/wpmu/sequences2012/files/2012/05/Buergin_LaCOSA_Poster.pdf

Burley, N. (1985). Leg-band color and mortality patterns in captive breeding populations of Zebra Finches. *The Auk, 102* (3), 647-651. doi: 82.43.98.187

Burnstein, D.D., & Wolff, P.C. (1967). Vocal conditioning in the guinea pig. *Psychonomic Science, 8* (1), 39-40. doi: 10.3758/BF03330656

Burt, J. M., Campbell, S. E., & Beecher, M. D. (2001). Song type matching as threat: a test using interactive playback. *Animal Behaviour, 62* (6), 1163-1170. doi: 10.1006/anbe.2001.1847

Byers, B. E., & Kroodsma, D. E. (2009). Female mate choice and songbird song repertoires. *Animal Behaviour, 77* (1), 13-22. doi: 10.1016/j.anbehav.2008.10.003

Byrne, R. W. (2002). Imitation of novel complex actions: what does the evidence from animals mean?. *Advances in the Study of Behavior, 31*, 77-105. doi: 10.1016/S0065-3454(02)80006-7

Caldwell, M. C. (1965). Individualized whistle contours in bottle-nosed dolphins (*Tursiops truncatus*). *Nature, 207*, 434-435. doi: 10.1038/207434a0

Caldwell, M. C., & Caldwell, D. K. (1979). The whistle of the Atlantic bottlenosed dolphin (*Tursiops truncatus*)—ontogeny. In: *Behavior of Marine Animals Vol 3, Cetaceans (*pp. 369-401). Springer, US.

Caldwell, M.C., Caldwell, D.K., & Tyack, P.L. (1990). Review of the signature-whistle-hypothesis for the Atlantic bottlenose dolphin. In: *The Bottlenose Dolphin* (pp. 199-234). Academic Press, US.

Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Convergence and divergence in Diana monkey vocalisations. *Biology Letters, 8* (3), 382-385. doi: 10.1098/rsbl.2011.1182

Candiotti, A., Zuberbühler, K., & Lemasson, A. (2013). Voice discrimination in four primates. *Behavioural Processes, 99*, 67-72. doi: 10.1016/j.beproc.2013.06.010

Carden, S. E., & Hofer, M. A. (1990). Independence of benzodiazepine and opiate action in the suppression of isolation distress in rat pups. *Behavioral Neuroscience, 104* (1), 160-166. doi: 10.1037/0735-7044.104.1.160

Carlson, B. E., & Langkilde, T. (2013). A common marking technique affects tadpole behavior and risk of predation. *Ethology*, *119* (2), 167-177. doi: 10.1111/eth.12050

Carter, R. T., Shaw, J. B., & Adams, R. A. (2014). Ontogeny of vocalisation in Jamaican fruit bats with implications for the evolution of echolocation. *Journal of Zoology*, *293* (1), 25-32. doi: 10.1111/jzo.12097

Cäsar, C., Zuberbühler, K., Young, R. J., & Byrne, R. W. (2013). Titi monkey call sequences vary with predator location and type. *Biology Letters*, *9* (5), 20130535. doi: 10.1098/rsbl.2013.0535

Casey, C., Reichmuth, C., Fregosi, S., Charrier, I., & Mathevon, N. (2013). Remember the titans: Individual vocal signatures reinforce dominance relationships among male northern elephant seals. *20th Biennial Conference on Marine Mammals*, Dunedin, New Zealand, 9-13 December.

Catchpole, C. K., & Slater, P. J. B. (1995). *Bird Song: Biological Themes and Variations* (pp. 253). Cambridge University, UK.

Catford, J.C. (2001) *A Practical Introduction to Phonetics* (pp. 248). Oxford University Press, US.

Caudron, A. K., Kondakov, A. A., & Siryanov, S. V. (1998). Acoustic structure and individual variation of grey seal (*Halichoerus grypus*) pup calls. *Journal of the Marine Biological Association of the United Kingdom*, *78* (2), 651-658. doi: 10.1017/S0025315400041680

Chan, A. A. Y. H., Giraldo-Perez, P., Smith, S., & Blumstein, D. T. (2010). Anthropogenic noise affects risk assessment and attention: the distracted prey hypothesis. *Biology Letters*, *6* (4), 458-461. doi: 10.1098/rsbl.2009.1081

Charlton, B. D., Reby, D., & McComb, K. (2007a). Female perception of size-related formant shifts in red deer, *Cervus elaphus*. *Animal Behaviour*, *74* (4), 707-714. doi: 10.1016/j.anbehav.2006.09.021

Charlton, B. D., Reby, D., & McComb, K. (2007b). Female red deer prefer the roars of larger males. *Biology Letters*, *3* (4), 382-385. doi: 10.1098/rsbl.2007.0244

Charlton, B. D., Reby, D., & McComb, K. (2008). Effect of combined source (F0) and filter (formant) variation on red deer hind responses to male roars. *The Journal of the Acoustical Society of America*, *123* (5), 2936-2943. doi: 10.1121/1.2896758

Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009). The information content of giant panda, *Ailuropoda melanoleuca*, bleats: acoustic cues to sex, age and size. *Animal Behaviour*, *78* (4), 893-898. doi: 10.1016/j.anbehav.2009.06.029

Charlton, B. D., & Reby, D. (2010). Context-related acoustic variation in male fallow deer (*Dama dama*) groans. *PloS One*, *6* (6), e21066-e21066. doi: 10.1371/journalpone.0021066

Charlton, B. D., Whisson, D. A., & Reby, D. (2012). Free-ranging male koalas use size-related variation in formant frequencies to assess rival males. *PloS One*, *8* (7), e70279-e70279. doi: 10.1371/journal.pone.0070279

Charlton, B. D., Frey, R., McKinnon, A. J., Fritsch, G., Fitch, W. T., & Reby, D. (2013). Koalas use a novel vocal organ to produce unusually low-pitched mating calls. *Current Biology*, *23* (23), R1035-R1036. doi: 10.1016/j.cub.2013.10.069

Charrier, I., Mathevon, N., & Jouventin, P. (2002). How does a fur seal mother recognize the voice of her pup? An experimental study of *Arctocephalus tropicalis*. *Journal of Experimental Biology*, *205* (5), 603-612.

Charrier, I., Mathevon, N., & Jouventin, P. (2003). Fur seal mothers memorize subsequent versions of developing pups' calls: adaptation to long-term recognition or evolutionary by-product? *Biological Journal of the Linnean Society*, *80* (2), 305-312. doi: 10.1046/j.1095-8312.2003.00239.x

Charrier, I., & Harcourt, R. G. (2006). Individual vocal identity in mother and pup Australian sea lions (*Neophoca cinerea*). *Journal of Mammalogy*, *87* (5), 929-938. doi: 10.1644/05-MAMM-A-344R3.1

Charrier, I., Pitcher, B. J., & Harcourt, R. G. (2009). Vocal recognition of mothers by Australian sea lion pups: individual signature and environmental constraints. *Animal Behaviour*, *78* (5), 1127-1134. doi: 10.1016/j.anbehav.2009.07.032

Charrier, I., Ahonen, H., & Harcourt, R.G. (2011). What makes an Australian seal lion (*Neophoca cinerea*) male's bark threatening? *Journal of Comparative Psychology, 125* (4)*, 385-392. doi: 10.1037/a0024513

Charrier, I. Mathevon, N., Casey, C., Fregosi, S., & Reichmuth, C. (2013). Size doesn't matter! Northern elephant seal males do not use phenotype cues in vocalisations to assess their rivals. *20th Biennial Conference on Marine Mammals*. Dunedin, New Zealand, 9-13 December.

Cheney, D. L., & Seyfarth, R. M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, *28* (2), 362-367. doi: 10.1016/S0003-3472(80)80044-3

Cheney, D.L., & Seyfarth, R.M. (1985). Vervet monkey alarm calls: manipulation through shared information? *Behaviour*, *94* (1), 150-166. doi: 10.1163/156853985X00316

Cheney, D. L., & Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, *36* (2), 477-486. doi: 10.1016/S0003-3472(88)80018-6

Chilvers, B. L., & Corkeron, P. J. (2001). Trawling and bottlenose dolphins' social structure. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268* (1479), 1901-1905. doi: 10.1098/rspb.2001.1732

Chomsky, N. (1957). Semantic structures (pp. 322). The Hague, Netherlands.

Clay, Z., Smith, C.L., & Blumstein, D.T. (2012). Food-associated vocalisations in mammals and birds: What do these calls really mean? *Animal Behavior, 8,* 323–330. doi: 10.1016/j.anbehav.2011.12.008

Clergeau, P., Croci, S., Jokimäki, J., Kaisanlahti-Jokimäki, M. L., & Dinetti, M. (2006). Avifauna homogenisation by urbanisation: analysis at different European latitudes. *Biological Conservation*, *127* (3), 336-344. doi: doi:10.1016/j.biocon.2005.06.035

Cohen, J. A., & Fox, M. W. (1976). Vocalisations in wild canids and possible effects of domestication. *Behavioural Processes*, *1* (1), 77-92. doi: 10.1016/0376-6357(76)90008-5

Collins, K. T., Terhune, J. M., Rogers, T. L., Wheatley, K. E. & Harcourt, R. G. (2006). Vocal individuality of in-air Weddell seal (*Leptonychotes weddellii*) pup 'primary' calls. *Marine Mammal Science*, *22* (4), 933-951. doi: 10.1111/j.1748-7692.2006.00074.x

Colvin, M. A. (1973). Analysis of acoustic structure and function in ultrasounds of neonatal Microtus. *Behaviour*, 234-263. doi: 82.29.174.250

Conner, D. A. (1982). Geographic variation in short calls of pikas (*Ochotona princeps*). *Journal of Mammalogy*, 48-52.

Cook, P., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology*, *127* (4), 412-427. doi: 10.1037/a0032345

Cooke, S. J., Iverson, S. J., Stokesbury, M. J., Hinch, S. G., Fisk, A. T., VanderZwaag, D. L., Apostle, R., & Whoriskey, F. (2011). Ocean Tracking Network Canada: a network approach to addressing critical issues in fisheries and resource management with implications for ocean governance. *Fisheries*, *36* (12), 583-592. doi: 10.1080/03632415.2011.633464

Corewyn, L. (2003). Ontogeny of Vocal Communication in Nonhuman Primates: A Review. *Laboratory Primate Newsletter*, *42* (1), 3-7.

Coscia, E. M., Phillips, D. P., & Fentress, J. C. (1991). Spectral analysis of neonatal wolf *Canis lupus* vocalisations. *Bioacoustics*, *3* (4), 275-293. doi: 10.1080/09524622.1991.9753190

Cosens, S. E. (1981). Development of vocalisations in the American Coot. *Canadian Journal of Zoology*, *59* (10), 1921-1928. doi: 10.1139/z81-262.

Costa, D. P., Le Boeuf, B. J., Huntley, A. C., & Ortiz, C. L. (1986). The energetics of lactation in the northern elephant seal, *Mirounga angustirostris*. *Journal of Zoology*, *209* (1), 21-33. doi: 10.1111/j.1469-7998.1986.tb03563.x

Crance, J. L., Bowles, A. E., & Garver, A. (2014). Evidence for vocal learning in juvenile male killer whales, *Orcinus orca*, from an adventitious cross-socializing experiment. *The Journal of Experimental Biology*, *217* (8), 1229-1237. doi: 10.1242/jeb.094300.

Cranford, T.W. (2000). In search of impulse sound sources in odontocetes. In: *Hearing by Whales and Dolphins* (pp. 109-155). Springer, US.

Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology*, *110* (3), 221-243. doi: 10.1111/j.1439-0310.2004.00968.x

Cunningham, K. A., Hayes, S. A., Rub, A. M. W., & Reichmuth, C. (2014). Auditory detection of ultrasonic coded transmitters by seals and sea lions. *The Journal of the Acoustical Society of America*, *135* (4), 1978-1985. doi: 10.11121/1.4868371

Curio, E. (1978). Adaptive significance of avian mobbing: Teleonomic hypotheses and predictions. *Zeitschrift für Tierpsychologie*, 48, 175-183. doi: 10.1111/j.1439-0310.1978.tb00255.x

Cuthill, I. C., Hunt, S., Cleary, C., & Clark, C. (1997). Colour bands, dominance, and body mass regulation in male zebra finches (*Taeniopygia guttata*). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *264* (1384), 1093-1099. doi: 10.1098/rspb.1997.0151

Darden, S. K., & Dabelsteen, T. (2006). Ontogeny of swift fox *Vulpes velox* vocalisations: production, usage and response. *Behaviour*, *143* (6), 659-682.

Darwin, C. (1871) *The Descent of Man, and Selection in Relation to Sex* (pp. 658). Library of Alexandria, UK.

Davies, C. E., Kovacs, K. M., Lydersen, C., & Van Parijs, S. M. (2006). Development of display behavior in young captive bearded seals. *Marine Mammal Science*, *22* (4), 952-965. doi: 10.1111/j/1748-7692.2006.00075.x

Davies, N. B. (2011). *Cuckoos, Cowbirds and Other Cheats* (pp. 312). A&C Black, US.

Dearolf, J. L., McLellan, W. A., Dillaman, R. M., Frierson, D., & Pabst, D. A. (2000). Precocial development of axial locomotor muscle in bottlenose dolphins (*Tursiops truncatus*). *Journal of Morphology*, *244* (3), 203-215. doi: 10.1002/(SICI)1097-4687(200006)244:3

Deecke, V. B., Ford, J. K., & Spong, P. (2000). Dialect change in resident killer whales: implications for vocal learning and cultural transmission. *Animal Behaviour*, *60* (5), 629-638. doi: 10.1006/anbe.2000.1454

Deecke, V. B., Slater, P. J., & Ford, J. K. (2002). Selective habituation shapes acoustic predator recognition in harbour seals. *Nature*, *420* (6912), 171-173. doi: 10.1038/nature01030

Dehnhardt, G., Mauck, B., Hanke, W., & Bleckmann, H. (2001). Hydrodynamic trail-following in harbor seals (*Phoca vitulina*). *Science*, *293* (5527), 102-104. doi: *10.1126/science.1060514*

DeRuiter, S. L., Boyd, I. L., Claridge, D. E., Clark, C. W., Gagnon, C., Southall, B. L., & Tyack, P. L. (2013). Delphinid whistle production and call matching during playback of simulated military sonar. *Marine Mammal Science*, *29* (2), E46-E59. doi: 10.1111/j.1748-7692.2012.00587.x

Ding, C., & He, X. (2004). K-means clustering via principle components analysis. In: *Proceedings of the Twenty-first International Conference on Machine Learning* (pp. 29). ACM, Canada.

Dobkin, D. S. (1979). Functional and evolutionary relationships of vocal copying phenomena in birds. *Zeitschrift für Tierpsychologie*, *50* (4), 348-363. doi: 10.1111/j.1439-0310.1979.tb01037.x

Eales, L. A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Animal Behaviour*, *37* (3), 507-508. doi: 10.1016/0003-3472(89)90097-3

Efremova, K. O., Volodin, I. A., Volodina, E. V., Frey, R., Lapshina, E. N., & Soldatova, N. V. (2011). Developmental changes of nasal and oral calls in the goitred gazelle, *Gazella subgutturosa*, a nonhuman mammal with a sexually dimorphic and descended larynx. *Naturwissenschaften, 98* (11), 919-931. doi: 10.1007/s00114-011-0843-7

Elowson, A. M., Snowdon, C. T., & Sweet, C. J. (1992). Ontogeny of trill and J-call vocalisations in the pygmy marmoset, *Cebuella pygmaea. Animal Behaviour, 43* (5), 703-715. doi: 10.1016/S0003-3472(05)80195-2

Elowson, A. M., & Snowdon, C. T. (1994). Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Animal Behaviour, 47* (6), 1267-1277. doi: 10.1006/anbe.1994.1175

Elowson, A. M., Snowdon, C. T., & Lazaro-Perea, C. (1998). Infant 'babbling' in a nonhuman primate: Complex vocal sequences with repeated call types. *Behaviour, 135* (5), 643-664. doi: 10.1163/156853998792897905

Elowson, A.M., Snowdon, C. T., & Lazaro-Perea, C. (1998). 'Babbling' and social context in infant monkeys: parallels to human infants. *Trends in Cognitive Sciences, 2* (1), 31-37. doi: 10.1016/S1364-6613(97)01115-7

Esser, K. H. (1994). Audio-vocal learning in a non-human mammal: the lesser spear-nosed bat, *Phyllostomus discolor. Neuroreport, 5* (14), 1718-1720.

Esser, K. H., & Schmidt, U. (1989). Mother-infant communication in the lesser Spear-nosed bat, *Phyllostomus discolor* (Chiroptera, Phyllostomidae)— Evidence for acoustic learning. *Ethology, 82* (2), 156-168. doi: 10.1111/j.1439-0310.1989.tb00496.x

Evans, W. E., Thomas, J. A., & Davis, R. W. (2004). Vocalisations from Weddell seals (*Leptonychotes weddellii*) during diving and foraging. In: *Echolocation in Bats and Dolphins* (pp. 541-547). The University of Chicago Press, US.

Ey, E., Hammerschmidt, K., Seyfarth, R. M., & Fischer, J. (2007). Age-and sex-related variations in clear calls of *Papio ursinus. International Journal of Primatology, 28* (4), 947-960. doi: 10.1007/s10764-007-9139-3

Fagot, J., & De Lillo, C. (2011). A comparative study of working memory: Immediate serial spatial recall in baboons (*Papio papio*) and humans. *Neuropsychologia, 49* (14), 3870-3880. doi: 10.1016/j.neuropsychologia.2011.10.003

Falls, J. B. (1992). Playback: a historical perspective. In: *Playback and Studies of Animal Communication* (pp. 11-33). Springer, US.

Falls, J. B., & Brooks, R. J. (1975). Individual recognition by song in white-throated sparrows. II. Effects of location. *Canadian Journal of Zoology*, *53* (10), 1412-1420. doi: 10.1139/z75-170

Fant, G. (1960). *Acoustic Theory of Speech Production* (pp. 328). Mouton and Co., Netherlands.

Farley, G.R., Barlow, S.M., & Netsell, R. (1992). Factors influencing neural activity in parabrachial regions during cat vocalisations. *Experimental Brain Research*, *89* (2), 341-351. doi: 10.1007/BF00228250

Favaro, L., Gnone, G., & Pessani, D. (2013). Postnatal development of echolocation abilities in a bottlenose dolphin (*Tursiops truncatus*): Temporal organization. *Zoo Biology*, *32* (2), 210-215. doi: 10.1002/zoo.21056

Fernández-Juricic, E., Campagna, C., Enriquez, V., & Ortiz, C. L. (1999). Vocal communication and individual variation in breeding South American sea lions. *Behaviour*, *136* (4), 495-517.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of Reinforcement* (pp. 724). Appleton Century Crofts, US.

Fichtel, C. (2004). Reciprocal recognition of sifaka (*Propithecus verreauxi verreauxi*) and redfronted lemur *(Eulemur fulvus rufus*) alarm calls. *Animal Cognition*, *7* (1), 45-52. doi: 10.1007/s10071-003-0180-0

Fisher, J. (1954). Evolution and bird sociality.In: *Evolution as a Process*, (pp. 71-83). Allen & Unwin, UK.

Fischer, J. (1998). Barbary macaques categorize shrill barks into two call types. *Animal Behaviour*, *55* (4), 799-807. doi: 10.1006/anbe.1997.0663

Fischer, J., Hammerschmidt, K., Cheney, D. L., & Seyfarth, R. M. (2002). Acoustic features of male baboon loud calls: influences of context, age, and individuality. *The Journal of the Acoustical Society of America*, *111* (3), 1465-1474. doi: 10.1121/1.1433807

Fiske, P., & Amundsen, T. (1997). Female bluethroats prefer males with symmetric colour bands. *Animal Behaviour*, *54* (1), 81-87. doi: 0003–3472/97/070081

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *The Journal of the Acoustical Society of America*, *102* (2), 1213-1222. doi: 10.1121/1.421048

Fitch, W. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, *4* (7), 258-267. doi: 10.1016/S1364-6613(00)01494-7

Fitch, W. T. (2005). The evolution of language: a comparative review. *Biology and Philosophy*, *20* (2-3), 193-203. doi: 10.1007/s10539-005-5597-1

Fitch, T. (2006). Production of vocalisations in mammals. *Visual Communication*, *3* (2), 145-175.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *106* (3), 1511-1522. doi: 10.1121/1.427148

Fitch, W. T., & Reby, D. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268* (1477), 1669-1675. doi: 10.1098/rspb.2001.1704

Fitch, W. T., & Hauser, M. D. (2003). Unpacking "honesty": vertebrate vocal production and the evolution of acoustic signals. In: *Acoustic Communication* (pp. 65-137). Springer, US.

Foote, A. D., Griffin, R. M., Howitt, D., Larsson, L., Miller, P. J., & Hoelzel, A. R. (2006). Killer whales are capable of vocal learning. *Biology Letters*, *2* (4), 509-512. doi: 10.1098/rsbl.2006.0525

Francis, C. D., Ortega, C. P., & Cruz, A. (2009). Noise pollution changes avian communities and species interactions. *Current Biology*, *19* (16), 1415-1419. doi: 10.1016/j.cub.2009.06.052

Freeberg, T.M. (2006) Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychological Science, 17,* 557-561.

Friedl, W.A., Nachtigall, P.E., Moore, P.W.B., Chun, N.K.W., Haun, J.E., Hall, R.W., & Richards, J.L. (1990). Taste reception in the bottlenose dolphin (*Tursiops truncatus gilli*) and the California sea lion (*Zalophus californianus*). In: *Sensory Abilities of Cetaceans: Laboratory and Field Evidence* (pp. 447-451). Plenum Press, US.

Fripp, D. (2005). Bubblestream whistles are not representative of a bottlenose dolphin's vocal repertoire. *Marine Mammal Science*, *21* (1), 29-44. doi: 10.1111/j.1748-7692.2005.tb01206.x

Fripp, D., Owen, C., Quintana-Rizzo, E., Shapiro, A., Buckstaff, K., Jankowski, K., Wells, R. & Tyack, P. (2005). Bottlenose dolphin (*Tursiops truncatus*) calves appear to model their signature whistles on the signature whistles of community members. *Animal Cognition*, *8* (1), 17-26. doi: 10.1007/s10071-004-0225-z

Frith, C., & Frith, D. (2004). *The Bowerbirds: Ptilonorhynchidae* (pp. 552). Oxford University Press, US.

Gabadinho, A., Ritschard, G., Müller, N. S., Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*, 40 (4), 1-37. http://www.jstatsoft.org/v40/i04/.

Garamszegi, L. Z., Pavlova, D. Z., Eens, M., & Møller, A. P. (2007). The evolution of song in female birds in Europe. *Behavioral Ecology*, *18* (1), 86-96. doi: 10.1093/beheco/arl047

Garcia, M., Charlton, B. D., Wyman, M. T., Fitch, W. T., & Reby, D. (2012). Do red deer stags (*Cervus elaphus*) use roar fundamental frequency (F0) to assess rivals?. *PloS One*, *8* (12), e83946-e83946. doi: 10.1371/journal.pone.00839946

Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, *165* (3894), 664-672. doi: 10.1126/science.165.3894.664

Garla, R. C., Chapman, D. D., Wetherbee, B. M., & Shivji, M. (2006). Movement patterns of young Caribbean reef sharks, *Carcharhinus perezi*, at Fernando de Noronha Archipelago, Brazil: the potential of marine protected areas for conservation of a nursery ground. *Marine Biology*, *149* (2), 189-199. doi: 10.1007/s00227-005-0201-4

Gaspard, J. C., Bauer, G. B., Reep, R. L., Dziuk, K., Cardwell, A., Read, L., & Mann, D.A. (2012). Audiogram and auditory critical ratios of two Florida manatees (*Trichechus manatus latirostris*). *The Journal of Experimental Biology*, *215* (9), 1442-1447. doi: 10.1242/jeb.065649

Gastwirth, J.L., Gel, Y.R., Hui, W.L.W., Lyubchich, V., Miao, W., & Noguchi, K., (2013). lawstat: An R package for biostatistics, public policy, and law. R package version 2.4.1.

Gellerman, L.W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *The Pedagogical Seminary and Journal of Genetic Psychology*, *42* (1), 206-208. doi:10.1080/08856559.1933.10534237

Ghazanfar, A. A., Flombaum, J. I., Miller, C. T., & Hauser, M. D. (2001). The units of perception in the antiphonal calling behavior of cotton-top tamarins (*Saguinus oedipus*): playback experiments with long calls. *Journal of Comparative Physiology A*, *187* (1), 27-35. doi: 10.1007/s003590000173

Gisiner, R., & Schusterman, R. J. (1991). California sea lion pups play an active role in reunions with their mothers. *Animal Behaviour*, *41* (2), 364-366. doi: 10.1016/S0003-3472(05)80488-9

Goodwin, S.E., & Podos, J. (2014). Team of rivals: alliance formations in territorial songbirds is predicted by vocal signal structure. *Biology Letters, 10* (2), 20131083. doi: 10.1098/rsbl.2013.1083

Götz, T. (2008). Aversiveness of sound in marine mammals: Psycho-physiological basis, behavioural correlates and potential applications. (pp. 221) Ph.D thesis, University of St. Andrews.

Götz, T., & Janik, V. M. (2010). Aversiveness of sounds in phocid seals: psycho-physiological factors, learning processes and motivation. *The Journal of Experimental Biology*, *213* (9), 1536-1548. doi: 10.1242/jeb.035535

Götz, T., & Janik, V. M. (2013). Acoustic deterrent devices to prevent pinniped depredation: efficiency, conservation concerns and possible solutions. *Marine Ecology Progress Series*, *492*, 285-302. doi: 10.3354/meps10482

Gould, E. (1975). Neonatal vocalisations in bats of eight genera. *Journal of Mammalogy*, 15-29. doi: 10.2307/1379603

Gould, E., Cooley, B., & Barnick, P. (1981). Echolocation by bats on their initial and subsequent flights. *Developmental Psychobiology, 14* (1), 41-54.

Grant, B. R., & Grant, P. R. (1996). Cultural inheritance of song and its role in the evolution of Darwin's finches. *Evolution*, *50* (6), 2471-2487.

Green, K., & Burton, H. R. (1988a). Annual and diurnal variations in the underwater vocalisations of Weddell seals. *Polar Biology*, *8* (3), 161-164. doi: 10.1007/BF00443447

Green, K., & Burton, H. R. (1988b). Do Weddell seals sing?. *Polar Biology*, *8* (3), 165-166. doi: 10.1007/BF00443448

Green, S. M. (1981). Sex differences and age gradations in vocalisations of Japanese and lion-tailed monkeys (*Macaca fuscata* and *Macaca silenus*). *American Zoologist*, *21* (1), 165-183. doi: 10.1093/icb/21.1.165

Greene, E., & Meagher, T. (1998). Red squirrels, *Tamiascuirus hudsonicus,* produce predator-class alarm calls. *Animal Behaviour*, *55* (3), 511-518. doi: 10.1006/anbe.1997.0620

Guinee, L. N., Chu, K., & Dorsey, E. M. (1983). Changes over time in the songs of known individual humpback whales (*Megaptera novaeangliae*). *Communication and Behavior of Whales*, *76*, 59-80.

Habersetzer, J., & Marimuthu, G. (1986). Ontogeny of sounds in the echolocating bat *Hipposideros speoris*. *Journal of Comparative Physiology A*, *158* (2), 247-257. doi: 10.1007/BF01338568

Hackett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 88-96.

Hagan, J.M., & Reed, J.M. (1988). Red colour bands reduce fledgling success in red-cockaded woodpeckers. *The Auk, 105* (3), 498-503. doi: 82.43.98.187

Hage, S. R., Gavrilov, N., & Nieder, A. (2013). Cognitive control of distinct vocalisations in rhesus monkeys. *Journal of Cognitive Neuroscience*, *25* (10), 1692-1701. doi: 10.1162/jocn_a_00428

Hammerschmidt, K., Newman, J. D., Champoux, M., & Suomi, S. J. (2000). Changes in rhesus macaque 'coo' vocalisations during early development. *Ethology*, *106* (10), 873-886. doi: 10.1046/j.1439-0310.2000.00611.x

Hammerschmidt, K., Freudenstein, T., & Jürgens, U. (2001). Vocal development in squirrel monkeys. *Behaviour*, *138*, 1179-1204. doi: 10.1016/S00033472(86)80252-4

Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. *The Evolution of Communicative Creativity: From Fixed Signals to Contextual Flexibility* (pp. 93-119). MIT Press Scholarship Online. doi: 10.7551/mitpress/9780262151214.003.0005

Hammill, M. O., & Gosselin, J. F. (1995). Grey seal (*Halichoerus grypus*) from the Northwest Atlantic: female reproductive rates, age at first birth, and age of maturity in males. *Canadian Journal of Fisheries and Aquatic Sciences*, *52* (12), 2757-2761. doi: 10.139/f95-864

Hanggi, E. B. (1992). The importance of vocal cues in mother–pup recognition in a California sea lion. *Marine Mammal Science*, *8* (4), 430-432. doi: 10.1111/j.1748-7692.1992.tb00061.x

Hanggi, E. B., & Schusterman, R. J. (1994). Underwater acoustic displays and individual variation in male harbour seals, *Phoca vitulina*. *Animal Behaviour*, *48* (6), 1275-1283. doi: 10.1006/anbe.1994.1363

Hanke, W., Brucker, C., & Bleckmann, H. (2000). The ageing of the low-frequency water disturbances caused by swimming goldfish and its possible relevance to prey detection. *Journal of Experimental Biology*, *203* (7), 1193-1200.

Hansen, P. (1979). Vocal learning: its role in adapting sound structures to long-distance propagation, and a hypothesis on its evolution. *Animal Behaviour*, *27*, 1270-1271. doi: 10.1016/0003-3472(79)90073-3

Hård, E., Engel, J., & Lindh, A. S. (1988). Effect of clonidine on ultrasonic vocalisation in preweaning rats. *Journal of Neural Transmission*, *73* (3), 217-237. doi: 10.1007/BF01250138

Harley, H. E., Crowell, S. E., Fellner, W., Odell, K., & Larsen-Plott, L. (2005). Rhythm perception and production by the bottlenose dolphin. *The Journal of the Acoustical Society of America*, *118* (3), 1906-1906. doi: 10.1121/1.4780115

Hasegawa, A., Okanoya, K., Hasegawa, T., & Seki, Y. (2011). Rhythmic synchronization tapping to an audio-visual metronome in budgerigars. *Scientific Reports*, *1*, 120. doi: 10.1038/srep00120

Hattori, Y., Tomonaga, M., & Matsuzawa, T. (2013). Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. *Scientific Reports*, *3,* 1566. doi: 10.1038/srep01566

Hauser, M.D. (1988). How infant vervet monkeys learn to recognize starling alarm calls: the role of experience. *Behaviour*, *105* (3), 187-201. doi: 10.1163/156853988X00016

Hauser, M. D. (1989). Ontogenetic changes in the comprehension and production of vervet monkey (*Cercopithecus aethiops*) vocalisations. *Journal of Comparative Psychology*, *103* (2), 149-158. doi: 10.1037/0735-7036.103.2.149

Hauser, M. D. (1992). Articulatory and social factors influence the acoustic structure of rhesus monkey vocalisations: a learned mode of production?. *The Journal of the Acoustical Society of America*, *91* (4), 2175-2179. doi: 10.1121/1.403676

Hauser, M. D. (1998). Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Animal Behaviour*, *55* (6), 1647-1658. doi: 10.1006/anbe.1997.0712

Hauser, M.D. & Wrangham, R.W. (1990). Recognition of predator and competitor calls in nonhuman primates and birds: a preliminary report. *Ethology, 86* (2), 116-130. doi: 10.1111/j.1439-0310.1990.tb00423.x

Hayes, S. A., Kumar, A., Costa, D. P., Mellinger, D. K., Harvey, J. T., Southall, B. L., & Le Boeuf, B. J. (2004). Evaluating the function of the male harbour seal, *Phoca vitulina*, roar through playback experiments. *Animal Behaviour*, *67* (6), 1133-1139. doi: 10.1016/j.anbehav.2003.06.019

Helverson, D.V. (2004). Object classification by echolocation in nectar feeding bats: size-independent generalization of shape. *Journal of Comparative Physiology A, 190* (7), 515-521. doi: 10.1007/s00359-004-0492-9

Herman, L. M., Pack, A. A., & Morrel-Samuels, P. (1993). Representational and conceptual skills of dolphins. In: *Comparative Cognition and Neuroscience* (pp. 403–442). Lawrence Erlbaum Associates, US.

Heupel, M. R., & Simpfendorfer, C. A. (2011). Estuarine nursery areas provide a low-mortality environment for young bull sharks *Carcharhinus leucas*. *Marine Ecology Progress Series*, *433*, 237-244. doi: 10.3354/meps09191

Hewer, H. R., & Backhouse, K. M. (1960). A preliminary account of a colony of grey seals, *Halichoerus grypus* (Fab.), in the southern Inner Hebrides. In *Proceedings of the Zoological Society of London*, *134* (2), 157-195. doi: 10.1111/j.1469-7998.1960.tb05586.x

Hiryu, S., & Riquimaroux, H. (2011). Developmental changes in ultrasonic vocalisations by infant Japanese echolocating bats, *Pipistrellus abramus*. *The Journal of the Acoustical Society of America*, *130* (4), EL147-EL153. doi: 10.1121/1.3632044

Hiss, A. (1983). Hoover. *New Yorker*; January 3, pp. 25–27.

Hoeschele, M., Cook, R. G., Guillette, L. M., Brooks, D. I., & Sturdy, C. B. (2012). Black-capped chickadee (*Poecile atricapillus*) and human (*Homo sapiens*) chord discrimination. *Journal of Comparative Psychology*, *126* (1), 57. doi: 10.1037/a0024627

Hofer, M. A. (1996). Multiple regulators of ultrasonic vocalisation in the infant rat. *Psychoneuroendocrinology*, *21* (2), 203-217. doi: 10.1016/0306-4530(95)00042-9

Hogstedt, G. (1983). Adaptation unto death: function of fear screams. *American Naturalist*, 562-570.

Hollén, L.I., & Manser, M.B. (2006). Ontogeny of alarm call responses in meerkats, *Suricata suricatta*: The roles of age, sex and nearby conspecifics. *Animal Behaviour*, *72*, 1345-1353. doi: 10.1016/j.anbehav.2006.03.020

Hollén, L. I., & Manser, M. B. (2007). Motivation before meaning: motivational information encoded in meerkat alarm calls develops earlier than referential information. *The American Naturalist*, *16 9*(6), 758-767.

Hollén, L. I., & Radford, A. N. (2009). The development of alarm call behaviour in mammals and birds. *Animal Behaviour*, *78* (4), 791-800. doi: 10.1016/j.anbehav.2009.07.021

Hunt, S., Cuthill, I. C., Swaddle, J. P., & Bennett, A. T. (1997). Ultraviolet vision and band-colour preferences in female zebra finches, *Taeniopygia guttata*. *Animal Behaviour*, *54* (6), 1383-1392. doi: 0003–3472/97/121383

Husson, F., Josse, J., Mazet, S.L., & Mazet J. (2014). FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R. R package version 1.25. http://CRAN.R-project.org/package=FactoMineR

Inoue, M. (1988). Age gradations in vocalisation and body weight in Japanese monkeys (*Macaca fuscata*). *Folia Primatologica*, *51* (2-3), 76-86. doi: 0015-5713/88/0513-0076$2.75/0

Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, *17* (23), R1004-R1005. doi: 10.1016/j.cub.2007.10.027

Insley, S. J. (1992). Mother-offspring separation and acoustic stereotypy: a comparison of call morphology in two species of pinnipeds. *Behaviour*, *120*, 103-122. doi: 10.1163/156853992X00237

Insley, S. J. (2000). Long-term vocal recognition in the northern fur seal. *Nature*, *406* (6794), 404-404. doi: 10.1038/35019064

Ivins, B. L., & Smith, A. T. (1983). Responses of pikas (*Ochotona princeps*, Lagomorpha) to naturally occurring terrestrial predators. *Behavioral Ecology and Sociobiology*, *13*(4), 277-285. doi: 10.1007/BF00299675

Janik, V. M. (1999). Origins and implications of vocal learning in bottlenose dolphins. In: *Mammalian Social Learning: Comparative and Ecological Perspectives* (pp. 308-326). Symposia of the Zoological society of London: Cambridge University Press, UK.

Janik, V. M. (2000). Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science*, *289* (5483), 1355-1357. doi: 10.1126/science.289.5483.1355

Janik, V. M. (2006). Communication in marine mammals. In *Encyclopedia of Language and Linguistics* (pp. 646-654). Elsevier, UK.

Janik, V., & Slater, P. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, *26*, 59-99. doi: 10.1016/S0065-3454(08)60377-0

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

Janik, V. M., & Slater, P. J. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, *60* (1), 1-11. doi: 10.1006/anbe.2000.1410

Janik, V. M., Sayigh, L. S., & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences*, *103* (21), 8293-8297. doi: 10.1073/pnas.0509918103

Jarvis, E.D. (2004). Learned birdsong and the neurobiology of human language. *Annals of the New York Academy of Sciences, 1016*, 749-777. 10.1196/anals.1298.038

Jarvis, E.D. (2006). Selection for and against vocal learning in birds and mammals. *Ornithological Science, 5*, 5-14. doi: 10.2326/osj.5.5

Jefferson, T. A., & Curry, B. E. (1996). Acoustic methods of reducing or eliminating marine mammal-fishery interactions: do they work?. *Ocean & Coastal Management*, *31* (1), 41-70. doi: 10.1016/0964-5691(95)00049-6

Job, D. A., Boness, D. J., & Francis, J. M. (1995). Individual variation in nursing vocalisations of Hawaiian monk seal pups, *Monachus schauinslandi* (Phocidae, Pinnipedia), and lack of maternal recognition. *Canadian Journal of Zoology*, *73* (5), 975-983. doi: 10.1139/z95-114

Jones, D. (1956). The pronunciation of English (pp. 229). Cambridge University Press, US.

Jones, G., & Ransome, R. D. (1993). Echolocation calls of bats are influenced by maternal effects and change over a lifetime. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *252* (1334), 125-128. 10.1098/rspb.1993.0055

Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: evidence for "fast mapping". *Science*, *304* (5677), 1682-1683. doi: 10.1126/science.1097859

Kastelein, R. A., Postma, J., & Verboom, W. C. (1995). Airborne vocalisations of Pacific walrus pups (*Odobenus rosmarus divergens*). In: *Sensory Systems of Aquatic Mammals* (pp. 265-285). De Spil Publishers, Netherlands.

Kawai, N., & Matsuzawa, T. (2000). Cognition: numerical memory span in a chimpanzee. *Nature*, *403* (6765), 39-40. doi: 10.1038/47405

Keesom, S.M., Rendon, N.M., Demas, G.E., & Hurley, L.M. (2015). Vocal behaviour during aggressive encounters between Siberian hamsters, *Phodopus sungorus*. *Animal Behaviour*, *102*, 85-93. doi: 10.1016/j.anbehav.2015.01.014

Kelley, L. A., Coe, R. L., Madden, J. R., & Healy, S. D. (2008). Vocal mimicry in songbirds. *Animal Behaviour*, *76* (3), 521-528. doi: 10.1016/j.anbehav.2008.04.012

Kerth, G., Perony, N., & Schweitzer, F. (2011). Bats are able to maintain long-term social relationships despite the high fission–fusion dynamics of their groups. *Proceedings of the Royal Society B: Biological Sciences*, *278* (1719), 2761-2767. doi:10.1098/rspb.2010.2718

Khan, C. B., Markowitz, H., & McCowan, B. (2006). Vocal development in captive harbor seal pups, *Phoca vitulina richardii*: Age, sex, and individual differences. *The Journal of the Acoustical Society of America*, *120* (3), 1684-1694. doi: 10.1121/1.2226530

Killebrew, D., Mercado, E., Herman, L., & Pack, A. (2001). Sound production of a neonate bottlenose dolphin. *Aquatic Mammals*, *27* (1), 34-44.

King, S. L., & Janik, V. M. (2013). Bottlenose dolphins can use learned vocal labels to address each other. *Proceedings of the National Academy of Sciences*, *110* (32), 13216-13221. doi: 10.1073/pnas.1304459110

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

King, S.L., & Janik, V.M. (2015). Come dine with me: food-associated social signalling in wild bottlenose dolphins (*Tursiops truncatus*). *Animal Cognition*, 1-6. doi: 10.1007/s10071-015-0851-7

Klenova, A. V., Goncharova, M. V., Bragina, E. V., & Kashentseva, T. A. (2014). Vocal development and voice breaking in Demoiselle Cranes (*Anthropoides virgo*). *Bioacoustics*, *23* (3), 1-19. doi: 10.1080/09524622.2014.900648

Knörnschild, M. (2014). Vocal production learning in bats. *Current Opinion in Neurobiology*, *28*, 80-85. doi: 10.1016/j.conb.2014.06.014

Knörnschild, M., Behr, O., & von Helversen, O. (2006). Babbling behavior in the sac-winged bat (*Saccopteryx bilineata*). *Naturwissenschaften*, *93* (9), 451-454. doi: 10.1007/s00114-006-0127-9

Knörnschild, M., Nagy, M., Metz, M., Mayer, F., & von Helversen, O. (2010). Complex vocal imitation during ontogeny in a bat. *Biology Letters*, *6* (2), 156-159. doi: 10.1098/rsbl.2009.0685

Koda, H., Oyakawa, C., Kato, A., & Masataka, N. (2007). Experimental evidence for the volitional control of vocal production in an immature gibbon. *Behaviour*, *144* (6), 681-692. doi: 10.1163/156853907781347817

Kooyman, G. L., & Kooyman, G. L. (1981). *Weddell Seal, Consummate Diver* (pp. 62-74). Cambridge University Press, US.

Koren, L. & Geffen, E. (2009). Complex call in male rock hyrax (*Procavia capensis*): a multi-information distributing channel. *Behavioural Ecology and Sociobiology, 63*, 581–590.

Kovacs, K. M., & Lavigne, D. M. (1986). Growth of grey seal (*Halichoerus grypus*) neonates: differential maternal investment in the sexes. *Canadian Journal of Zoology*, *64* (9), 1937-1943. doi: 10.1139/z86-291

Kroodsma, D. E. (1982). Learning and the ontogeny of sound signals in birds. *Acoustic Communication in Birds*, *2*, 1-23.

Kuhl, P. K. (1981). Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *The Journal of the Acoustical Society of America*, *70* (2), 340-349. doi: 10.1121/1.386782

Lachlan, R. F., & Slater, P. J. (1999). The maintenance of vocal learning by gene–culture interaction: the cultural trap hypothesis. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266* (1420), 701-706. doi: 10.1098/rspb.1999.0692

Lade, B. I., & Thorpe, W. H. (1964). Dove songs as innately coded patterns of specific behaviour. *Nature*, *202*, 366-368. doi: 10.1038/202366a0

Lal, H. (1967). Operant control of vocal responding in rats. *Psychonomic Science*, *8* (1), 35-36. doi: 10.3758/BF03330654

Lapshina, E. N., Volodin, I. A., Volodina, E. V., Frey, R., Efremova, K. O., & Soldatova, N. V. (2012). The ontogeny of acoustic individuality in the nasal calls of captive goitred gazelles, *Gazella subgutturosa*. *Behavioural Processes*, *90* (3), 323-330. doi: 10.1016/j.beproc.2012.03.011

Le Prell, C.G., Hauser, M.D., & Moody, D.B. (2002). Discrete or graded variation within rhesus monkey screams? Psychophysical experiments on classification. *Animal Behaviour, 63* (1), 47-62. doi: 10.1006/anbe.2001.1888

Lent, K. (1989). An efficient method for pitch shifting digitally sampled sounds. *Computer Music Journal*, 65-71. doi: 10.2307/3679554

Lenth, R.V. (2013). lsmeans: Least-squares means. R package version 1.10-01. http://CRAN.R-project.org/package=lsmeans

Le Roux, A., Cherry, M.I., & Manser, M.B. (2009). The vocal repertoire in a solitary foraging carnivore, *Cynictis penicillate*, may reflect facultative sociality. *Naturwissenschaften, 96,* 575-584.

Lesage, V., Barrette, C., Kingsley, M., & Sjare, B. (1999). The effect of vessel noise on the vocal behavior of belugas in the St. Lawrence River estuary, Canada. *Marine Mammal Science*, *15* (1), 65-84. doi: 10.1111/j.1748-7692.1999.tb00782.x

Li, H. S., & Borg, E. (1991). Age-related loss of auditory sensitivity in two mouse genotypes. *Acta Oto-laryngologica*, *111* (3), 827-834.

Li, H.S., Wang, D., Wang, K., Xiao, J., & Akamatsu, T. (2007). The ontogeny of echolocation in a Yangtze finless porpoise (*Neophocaena phocaenoides asiaeorientalis*). *The Journal of the Acoustical Society of America*, *122* (2), 715-718. doi:10.1121/1.2747203

Lidgard, D. C., Bowen, W. D., Jonsen, I. D., & Iverson, S. J. (2012). Animal-borne acoustic transceivers reveal patterns of at-sea associations in an upper-trophic level predator. *PloS one*, *7* (11), e48962. doi:10.1371/journal.pone.0048962

Lilly, J. C. (1965). Vocal mimicry in Tursiops: Ability to match numbers and durations of human vocal bursts. *Science*, *147* (3655), 300-301. doi: 10.1126/science.147.3655.300

Lind, H., Dabelsteen, T., & McGregor, P. K. (1996). Female great tits can identify mates by song. *Animal Behaviour*, *52* (4), 667-671. doi: 10.1006/anbe.1996.0211

Lindhard, M. (1988) Apparent sonar clicks from a captive bottlenose dolphin. In: *Animal Sonar* (pp. 109-113). Springer, US.

Locke, J.L. (1986). The linguistic significance of babbling. In*: Prescursors of Early Speech* (pp. 143-160). Stockton Press, Netherlands.

Loffredo, C. A., & Borgia, G. (1986). Male courtship vocalisations as cues for mate choice in the satin bowerbird. *The Auk*, *103*, 189-195.

Lusseau, D., Schneider, K., Boisseau, O. J., Haase, P., Slooten, E., & Dawson, S. M. (2003). The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, *54* (4), 396-405. doi:10.1007/s00265-003-0651-y

Macedonia, J. M. (1993). The vocal repertoire of the ringtailed lemur (*Lemur catta*). *Folia Primatologica*, *61* (4), 186-217. doi: 0614-0186$8.00.0

Macphail, E.M., (1982). *Brain and Intelligence in Vertebrates* (pp. 398). Oxford, UK.

Madsen, P. T., Carder, D. A., Au, W. W. L., Nachtigall, P. E., Møhl, B., & Ridgway, S. H. (2003). Sound production in neonate sperm whales (L). *The Journal of the Acoustical Society of America*, *113* (6), 2988-2991. doi: 10.1121/1.1572137

Madsen, P. T., Jensen, F. H., Carder, D., & Ridgway, S. (2012). Dolphin whistles: a functional misnomer revealed by heliox breathing. *Biology Letters*, *8* (2), 211-213. doi: 10.1098/sbl.2011.0701

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2014). cluster: Cluster analysis basics and extensions. R package version 1.15.2

Malinka, C. (2014). Baby's first sonar: development of echolocation in the bottlenose dolphin (*Tursiops truncatus*). Master's thesis, University of St. Andrews. 44 pp.

Mandelli, M. J., & Sales, G. (2004). Ultrasonic vocalisations of infant short-tailed field voles, *Microtus agrestis*. *Journal of Mammalogy*, *85* (2), 282-289. doi: 10.1644/1545-1542(2004)085<0282:UVOISF>2.0.CO;2

Manly, B.F.J. (1997) Randomization, bootstrap and Monte Carlo methods in biology (pp.461). Chapman and Hall, UK.

Manoukian, S., Azzali, M., Farchi, C., & Tizzi, R. (2002). Development of *Tursiops truncatus* sonar system during first months of life in captivity: acoustic interactions with other community members. *Acta Acustica United with Acustica*, *88* (5), 743-747.

Manser, M.B., Seyfarth, R.M., & Cheney, D.L. (2002). Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Sciences*, *6* (2), 55-57. Doi: 10.1016/S1364-6613(00)01840-4

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27* (2), 209-220.

Marler, P., & Peters, S. (1982). Subsong and plastic song: their role in the vocal learning process. In: *Acoustic Communication in Birds* (pp. 25-50). Academic press, US.

Marler, P., & Nelson, D. A. (1992). Action-based learning: a new form of developmental plasticity in bird song. *Netherlands Journal of Zoology*, *43* (1), 91-103. doi: 10.1163/156854293X00232

Maros, K., Pongrácz, P., Bárdos, G., Molnár, C., Faragó, T., & Miklósi, Á. (2008). Dogs can discriminate barks from different situations. *Applied Animal Behaviour Science*, *114* (1), 159-167. doi: 10.1016/j.applanim.2008.01.022

Marshall, A. J., Wrangham, R. W., & Arcadi, A. C. (1999). Does learning affect the structure of vocalisations in chimpanzees?. *Animal Behaviour, 58* (4), 825-830. doi: 10.1006/anbe.1999.1219

Masataka, N., & Fujita, K. (1989). Vocal learning of Japanese and rhesus monkeys. *Behaviour, 109* (3-4), 191-199. doi: 10.1163/156853989X00222

Mateo, J. (1996a). Early auditory experience and the ontogeny of alarm call discrimination in Belding's ground squirrels (*Spermophilus beldingi*). *Journal of Comparative Psychology, 110* (2), 115–124. doi: 10.1037/0735-7036.110.2.115

Mateo, J. (1996b). The development of alarm call response behavior in free-living juvenile Belding's ground squirrels. *Animal Behaviour, 52* (3), 489–505. doi: 10.1006/anbe.1996.0192

Mateo, J.M. (1999). Plasticity of alarm-call response development in Belding's ground squirrels (*Spermophilus beldingi,* Sciuridae). *Ethology, 105,* 193-206.

Matrosova, V. A., Volodin, I. A., Volodina, E. V., & Babitsky, A. F. (2007). Pups crying bass: vocal adaptation for avoidance of age-dependent predation risk in ground squirrels?. *Behavioral Ecology and Sociobiology, 62* (2), 181-191. doi: 10.1007/s00265-007-0452-9

Matsumura, S. (1979). Mother-infant communication in a horseshoe bat (*Rhinolophus ferrumequinum nippon*): development of vocalisation. *Journal of Mammalogy, 60* (1), 76-84.

McComb, K., & Semple, S. (205). Co-evolution of vocal communication and sociality in primates. *Biology Letters, 1,* 381-385.

McCowan, B., & Reiss, D. (1995). Whistle contour development in captive-born infant bottlenose dolphins (*Tursiops truncatus*): Role of learning. *Journal of Comparative Psychology, 109* (3), 242. doi: 10.1037/0735-7036.109.3.242

McCulloch, S. (1999). The vocal behaviour of the grey seal (*Halichoerus grypus*) (pp.150). Ph.D. thesis, University of St Andrews.

McCulloch, S., Pomeroy, P. P., & Slater, P. J. (1999). Individually distinctive pup vocalisations fail to prevent allo-suckling in grey seals. *Canadian Journal of Zoology, 77* (5), 716-723. doi: 10.1139/z99-023

McCulloch, S., & Boness, D. J. (2000). Mother–pup vocal recognition in the grey seal (*Halichoerus grypus*) of Sable Island, Nova Scotia, Canada. *Journal of Zoology, 251* (4), 449-455. doi: 10.1111/j.1469-7998.2000.tb00800.x

Mellish, J. A., Iverson, S. J., & Don, B. W. (1999). Variation in milk production and lactation performance in grey seals and consequences for pup growth and weaning characteristics. *Physiological and Biochemical Zoology*, *72* (6), 677. doi: 10.1086/316708

Melnychuk, M. C., Welch, D. W., & Walters, C. J. (2010). Spatio-temporal migration patterns of Pacific salmon smolts in rivers and coastal marine waters. *PLoS One*, *5* (9), e12916. doi:10.1371/journal.pone.0012916

Mercado III, E., Killebrew, D.A., Pack, A.A., Mácha, I.V.B., & Herman, L.M. (2000). Generalization of 'same-different' classification abilities in bottlenosed dolphins. *Behavioural Processes*, *50* (2), 79-94. doi: 10.1016/S0376-6357(00)00100-5

Mercado III, E., Schneider, J. N., Pack, A. A., & Herman, L. M. (2010). Sound production by singing humpback whales. *The Journal of the Acoustical Society of America*, *127* (4), 2678-2691. doi: 10.1121/1.3309453

Miksis, J. L., Tyack, P. L., & Buck, J. R. (2002). Captive dolphins, *Tursiops truncatus*, develop signature whistles that match acoustic features of human-made model sounds. *The Journal of the Acoustical Society of America*, *112* (2), 728-739. doi: 10.1121/1.1496079

Miller, E. H., & Murray, A. V. (1995). Structure, complexity, and organization of vocalisations in harp seal (Phoca groenlandica) pups. In: *Sensory Systems of Aquatic Mammals (*pp. 237-264). DeSpil Publishers, Netherlands.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, *63* (2), 81-97.

Mitani, J. C., Hasegawa, T., Gros-Louis, J., Marler, P., & Byrne, R. (1992). Dialects in wild chimpanzees?. *American Journal of Primatology*, *27* (4), 233-243. doi: 10.1002/ajp.1350270402

Mitani, J., & Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour*, *135* (8-9), 1041-1064. doi: 10.1163/156853998792913483

Molliver, M.E. (1963). Operant control of vocal behavior in the cat. *Journal of the Experimental Analysis of Behavior*, *6* (2), 197-202. doi: 10.1901/jeab.1963.6-197

Monroy, J. A., Carter, M. E., Miller, K. E., & Covey, E. (2011). Development of echolocation and communication vocalisations in the big brown bat, *Eptesicus fuscus. Journal of Comparative Physiology A*, *197* (5), 459-467. doi: 10.1007/s00359-010-0614-5

Moore, B. R. (1992). Avian movement imitation and a new form of mimicry: tracing the evolution of a complex form of learning. *Behaviour*, *122* (3), 231-263. doi: 10.1163/156853992X00525

Moore, B.R. (1996). The evolution of imitative learning. In: *Social Learning in Animals: The Roots of Culture* (pp. 245-265). Academic Press, US.

Morisaka, T., Shinohara, M., & Taki, M. (2005). Underwater sounds produced by neonatal bottlenose dolphins (*Tursiops truncatus*): I. Acoustic characteristics. *Aquatic Mammals*, *31* (2), 248-257. doi: 10.1578/AM.31.2.2005.248

Morrice, M. G., Burton, H. R., & Green, K. (1994). Microgeographic variation and songs in the underwater vocalisation repertoire of the Weddell seal (*Leptonychotes weddellii*) from the Vestfold Hills, Antarctica. *Polar Biology*, *14* (7), 441-446. doi: 10.1007/BF00239046

Morton, E. (1977). On the occurrence and significance of motivation-structural rules in some birds and mammal sounds. *American Naturalist, 111*, 855–869.

Morton, E. S. (1996). Why songbirds learn songs: An arms race over ranging?. *Poultry and Avian Biology Reviews*, *7*, 65-71.

Moss, C.F. (1988). Ontogeny of vocal signals in the big brown bat, *Eptesicus fuscus*. In: *Animal Sonar* (pp. 115-120). Springer, US.

Moss, C. F., Redish, D., Gounden, C., & Kunz, T. H. (1997). Ontogeny of vocal signals in the little brown bat, *Myotis lucifugus*. *Animal Behaviour*, *54* (1), 131-141. doi: 10.1006/anbe.1996.0410

Motomura, N., Shimizu, K., Shimizu, M., Aoki-Komori, S., Taniguchi, K., Serizawa, I., & Saito, T. R. (2002). A comparative study of isolation-induced ultrasonic vocalisation in rodent pups. *Experimental Animals*, *51* (2), 187-190. doi: 10.1538/expanim.51.187

Muñoz, G., Karamanlidis, A. A., Dendrinos, P., & Thomas, J. A. (2011). Aerial vocalisations by wild and rehabilitating Mediterranean monk seals (*Monachus monachus*) in Greece. *Aquatic Mammals*, *37* (3), 262. doi: 10.1578/AM.37.3.3011.262

Musser, W. B., Bowles, A. E., Grebner, D. M., & Crance, J. L. (2014). Differences in acoustic features of vocalisations produced by killer whales cross-socialized with bottlenose dolphins. *The Journal of the Acoustical Society of America*, *136* (4), 1990-2002. doi: 10.1121/1.4893906

Myers, S. A., Horel, J. A., & Pennypacker, H. S. (1965). Operant control of vocal behavior in the monkey *Cebus albifrons*. *Psychonomic Science*, *3* (1-12), 389-390. doi: 10.3758/BF03343194

Nachtigall, P. E., Yuen, M. M., Mooney, T. A., & Taylor, K. A. (2005). Hearing measurements from a stranded infant Risso's dolphin, *Grampus griseus*. *Journal of Experimental Biology*, *208* (21), 4181-4188. doi: 10.1242/jeb.01876

Nachtigall, P. E., Mooney, T. A., Taylor, K. A., & Yuen, M. M. (2007). Hearing and auditory evoked potential methods applied to odontocete cetaceans. *Aquatic Mammals*, *33* (1), 6. doi: 10.1578/AM.33.1.2007.6

Neiworth, J.J., & Wright, A.A. (1994). Monkeys (*Macaca mulatta*) learn category matching in a nonidentical same-different task. *Journal of Experimental Psychology: Animal Behaviour Processes*, *20* (4), 429-435. doi: 10.1037/0097-7403.20.4.429

Nelson, J. E. (1965). Behaviour of Australian pteropodidae (Megacheroptera). *Animal Behaviour*, *13* (4), 544-557. doi: 10.1016/0003-3472(65)90118-1

Noad, M. J., Cato, D. H., Bryden, M. M., Jenner, M. N., & Jenner, K. C. S. (2000). Cultural revolution in whale songs. *Nature*, *408* (6812), 537-537. doi: 10.1038/35046199

Noirot, E., & Pye, D. (1969). Sound analysis of ultrasonic distress calls of mouse pups as a function of their age. *Animal Behaviour*, *17*, 340-349. doi: 10.1016/0003-3472(69)90020-7

Nordby, J. C., Campbell, S. E., & Beecher, M. D. (2001). Late song learning in song sparrows. *Animal Behaviour*, *61* (4), 835-846. doi: 10.1006/anbe.2000.1673

Nottebohm, F. (1970). Ontogeny of bird song. *Science*, *167* (3920), 950-956. doi: 10.1126/science.167.3920.950

Nottebohm, F., & Nottebohm, M. E. (1971). Vocalisations and breeding behaviour of surgically deafened ring doves (*Streptopelia risoria*). *Animal Behaviour*, *19* (2), 313-327. doi: 10.1016/S0003-3472(71)80012-X

Nottebohm, F. (1972). The origins of vocal learning. *American Naturalist*, *106*, 116-140.

Nottebohm, F. (1991). Reassessing the mechanisms and origins of vocal learning in birds. *Trends in Neurosciences*, *14* (5), 206-211. doi: 10.1016/0166-2236(91)90107-6

Nowacek, D. P., Thorne, L. H., Johnston, D. W., & Tyack, P. L. (2007). Responses of cetaceans to anthropogenic noise. *Mammal Review*, *37* (2), 81-115. doi: 10.1111/j.1365-2907.2007.00104.x

Nowicki, S., Peters, S., & Podos, J. (1998). Song learning, early nutrition and sexual selection in songbirds. *American Zoologist*, *38* (1), 179-190. doi: 10.1093/icb/38.1.179

Nowicki, S., Hasselquist, D., Bensch, S., & Peters, S. (2000). Nestling growth and song repertoire size in great reed warblers: evidence for song learning as an indicator mechanism in mate choice. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *267* (1460), 2419-2424. doi: 10.1098/rspb.2000.1300

Nowicki, S., Searcy, W. A., & Peters, S. (2002). Quality of song learning affects female response to male bird song. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269* (1503), 1949-1954. doi: 10.1098/rspb.2002.2124

Nowicki, S., & Searcy, W. A. (2005). Song and mate choice in birds: how the development of behavior helps us understand function. *The Auk*, *122* (1), 1-14. doi: 10.1642/0004-8038(2005)122[0001:SAMCIB]2.0.CO;2

O'Loghlen, A. L., & Rothstein, S. I. (2012). When less is best: female brown-headed cowbirds prefer less intense male displays. *PloS one*, *7* (5), e36130. doi: 10.1371/journal.pone.0036130

Odom, K. J., Hall, M. L., Riebel, K., Omland, K. E., & Langmore, N. E. (2013). Female song is widespread and ancestral in songbirds. *Nature Communications*, *5*, 3379-3379. doi: 10.1038/ncomms4379

Oetelaar, M. L., Terhune, J. M., & Burton, H. R. (2003). Can the sex of a Weddell seal (*Leptonychotes weddellii*) be identified by its surface call?. *Aquatic Mammals*, *29* (2), 261-267.

Oksanen, J., Guillaume Blanchet, F., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., & Wagner, H. (2013). vegan: Community ecology package. R package version 2.0-10.

Oliver, G. W. (1978). Navigation in mazes by a grey seal, *Halichoerus grypus* (Fabricus). *Behaviour*, *67* (1-2), 97-114. doi: 10.1163/156853978X00279

Oller, D.K. (1980) The emergence of the sounds of speech in infancy. In *Child Phonology: Production* (pp. 93-112). New York: Academic Press.

Oller, D. K., Wieman, L. A., Doyle, W. J., & Ross, C. (1976). Infant babbling and speech. *Journal of Child Language, 3* (1), 1-11. doi: 10.1017/S0305000900001276

Oller, D.K., & Eilers, R.E. (1988). The role of audition in infant babbling. *Child Development, 59* (2), 441-449. doi: 10.2307/1130323

Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (1993). Vocalisations of rhesus (*Macaca mulatta*) and Japanese (*M. Fuscata*) macaques cross-fostered between species show evidence of only limited modification. *Developmental Psychobiology, 26* (7), 389-406. doi: 10.1002/dev/420260703

Owren, M.J., Seyfarth, R.M., & Cheney, D.L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (*Papio cyncephalus ursinus*): implications for production processes and functions. *The Journal of the Acoustical Society of America, 101* (5), 2951-2963. doi: 10.1121/1.418523

Padilla de la Torre, M., Briefer, E.F., Reader, T., & McElligott, A.G. (2015). Acoustic analysis of cattle (*Bos taurus*) mother-offspring contact calls from a source-filter theory perspective. *Applied Animal Behaviour Science, 163,* 58-68. doi: 10.1016/j.applanim.2014.11.017

Page, B., Goldsworthy, S. D., & Hindell, M. A. (2002). Individual vocal traits of mother and pup fur seals. *Bioacoustics, 13* (2), 121-143. doi: 10.1080/09524622.2002.9753491

Page, R.A., & Ryan, M.J. (2006). Social transmission of novel foraging behaviour in bats: Frog calls and their referents. *Current Biology, 16,* 1201-1205. doi: 10.1016/j.cub/2006/04/038

Patel, A. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception, 24,* 99-104. doi: 10.1525/mp.2006.24.1.99

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology, 19* (10), 827-830. doi: 10.1016/j.cub.2009.03.038

Paterson, W. D., Redman, P., Hiby, L. A., Moss, S. E., Hall, A. J., & Pomeroy, P. (2013). Pup to adult photo-ID: Evidence of pelage stability in gray seals. *Marine Mammal Science, 29* (4), E537-E541. doi: 10.1111/mms.12043

Patterson, F. G. (1978). The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language, 5* (1), 72-97. doi: 10.1016/0093-934X(78)90008-1

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

Payne, R. B. (1982). Ecological consequences of song matching: Breeding success and intraspecific song mimicry in indigo buntings. *Ecology*, *63* (2), 401-411. doi: 10.2307/1938958

Payne, R. B. (1983). The social context of song mimicry: Song-matching dialects in indigo buntings (*Passerina cyanea*). *Animal Behaviour*, *31* (3), 788-805. doi: 10.1016/S0003-3472(83)80236-X

Pearce, J.M. (1987). Communication and language. In: *Introduction to Animal Cognition* (pp. 251-283). Lawrence Erlbaum Associates, US.

Pepperberg, I. M. (1994). Vocal learning in grey parrots (*Psittacus erithacus*): effects of social interaction, reference, and context. *The Auk*, *111* (2), 300-313. doi: 10.2307/4088595

Pepperberg, I.M. (1996). Categorical class formation by an African grey parrot (*Psittacus erithacus*). In: *Advances in Psychology: Stimulus Class Formation in Humans and Animals* (pp. 71-90). Elsevier, Netherlands. doi: 10.1016/S0166-4115(06)80104-4

Pepperberg, I.M. (1999). *The Alex Studies: Cognitive and communicative abilities of grey parrots* (pp. 448). Harvard University Press, US.

Pepperberg, I.M., & Brezinsky, M.V. (1991). Acquisition of a relative class concept by an African gray parrot (*Psittacus erithacus*): Discrimination based on relative size. *Journal of Comparative Psychology*, *105* (3), 286-294. doi: 10.1037/0735-7036.105.3.286

Peters, G. (2011). Dominant frequency of loud mew calls of felids (Mammalia: Carnivora) decreases during ontogenetic growth. *Mammal Review*, *41* (1), 54-74. doi: 10.111/j.1365-2907.2010.00169.x

Pfungst, O. (1911). *Clever Hans: (the horse of Mr. Von Osten.) A Contribution to Experimental Animal and Human Psychology* (pp. 112). Holt, Rinehart and Winston, US.

Pistorio, A. L., Vintch, B., & Wang, X. (2006). Acoustic analysis of vocal development in a New World primate, the common marmoset (*Callithrix jacchus*). *The Journal of the Acoustical Society of America*, *120* (3), 1655-1670. doi: 10.1121/1.2225899

Pitcher, B. J., Ahonen, H., Harcourt, R. G., & Charrier, I. (2009). Delayed onset of vocal recognition in Australian sea lion pups (*Neophoca cinerea*). *Naturwissenschaften*, *96* (8), 901-909. doi: 10.1007/s00114-009-0546-5

Podos, J. (1996). Motor constraints on vocal development in a songbird. *Animal Behaviour*, *51* (5), 1061-1070. doi: 10.1006/anbe.1996.0107

Pomeroy, P. P., Anderson, S. S., Twiss, S. D., & McConnell, B. J. (1994). Dispersion and site fidelity of breeding female grey seals (*Halichoerus grypus*) on North Rona, Scotland. *Journal of Zoology*, *233* (3), 429-447. doi: 10.1111/j.1469-7998.1994.tb05275.x

Pomeroy, P. P., Twiss, S. D., & Duck, C. D. (2000). Expansion of a grey seal (*Halichoerus grypus*) breeding colony: changes in pupping site use at the Isle of May, Scotland. *Journal of Zoology*, *250* (1), 1-12. doi: 10.1111/j.1469-7998.2000.tb00573.x

Pomeroy, P. P., Redman, P. R., Ruddell, S. J. S., Duck, C. D., & Twiss, S. D. (2005). Breeding site choice fails to explain interannual associations of female grey seals. *Behavioral Ecology and Sociobiology*, *57* (6), 546-556. doi: 10.1007/s00265-004-0882-6

Pontet, A., Gyger, M., & Schenk, F. (1989). Ontogeny of ultrasonic vocalisations in the woodmouse (*Apodemus sylvaticus L.*). I: Temporal organization. *Behaviour*, *108* (3-4), 3-4. doi: 82.29.174.250

Poole, J. H., Tyack, P. L., Stoeger-Horwath, A. S., & Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. *Nature*, *434* (7032), 455-456. doi: 10.1038/434455a

Porter, D., & Neuringer, A. (1984). Music discriminations by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *10* (2), 138. doi: 10.1037/0097-7403.10.2.138

Premack, A. J., & Premack, D. (1972). Teaching language to an ape. *Scientific American*, *227* (4), 92-99.

Quick, N.J. & Janik, V.M. (2012). Bottlenose dolphins exchange signature whistles when meeting at sea. *Proceedings of the Royal Society B: Biological Sciences*, *279* (1738), 2539-2545. doi: 10.1098/rspb.2011.2537

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rainey, H.J., Zuberbühler, K., & Slater, P.J.B. (2004a). Hornbills can distinguish between primate alarm calls. *Proceedings of the Royal Society B: Biological Sciences, 271* (1540), 755-759. doi: 10.1098/rspb.2003.2619

Rainey, H.J., Zuberbühler, K., & Slater, P.J.B. (2004b). The responses of black-casqued hornbills to predator vocalisations and primate alarm calls. *Behaviour*, *141* (10), 1263-1277. doi: 10.1163/1568539042729658

Ralls, K., Fiorelli, P., & Gish, S. (1985). Vocalisations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Canadian Journal of Zoology*, *63* (5), 1050-1056. doi: 10.1139/z85-157

Randolph, M. C., & Brooks, B. A. (1967). Conditioning of a vocal response in a chimpanzee through social reinforcement. *Folia Primatologica*, *5* (1-2), 70-79. doi: 10.1159/000161938

Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs. *Animal Cognition*, *11* (2), 339-347. doi: 10.1007/s10071-007-0123-2

Reby, D., Cargnelutti, B., Joachim, J., & Aulagnier, S. (1999). Spectral acoustic structure of barking in roe deer (*Capreolus capreolus*). Sex-, age-and individual-related variations. *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie*, *322* (4), 271-279. doi: 10.1016/S0764-4469(99)80063-8

Reby, D., & McComb, K. (2003). Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour*, *65* (3), 519-530. doi: 10.1006/anbe.2003.2078

Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., & Clutton-Brock, T. (2005). Red deer stags use formants as assessment cues during intrasexual agonistic interactions. *Proceedings of the Royal Society B: Biological Sciences*, *272* (1566), 941-947. doi: 10.1098/rspb.2004.2954

Reby, D., Charlton, B. D., Locatelli, Y., & McComb, K. (2010). Oestrous red deer hinds prefer male roars with higher fundamental frequencies. *Proceedings of the Royal Society of London B: Biological Sciences*, *277* (1695), 2747-2753. doi: 10.1098/rspb.2010.0467

Redman, P. 2002. The role of temporal, spatial and kin associations in grey seal breeding colonies (pp. 245). Ph.D. thesis, University of St. Andrews.

Reichmuth, C., & Casey, C. (2014). Vocal learning in seals, sea lions, and walruses. *Current Opinion in Neurobiology*, *28*, 66-71. doi: 10.1016/j.conb.2014.06.011

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

Reiss, D. (1988) Observations on the development of echolocation in young bottlenose dolphins. In: *Animal Sonar (*pp. 121-127). Springer, US.

Reiss, D., & McCowan, B. (1993). Spontaneous vocal mimicry and production by bottlenose dolphins (*Tursiops truncatus*): Evidence for vocal learning. *Journal of Comparative Psychology*, *107* (3), 301. doi: 10.1037/0735-7036.107.3.301

Renouf, D. (1984). The vocalisation of the harbour seal pup (*Phoca vitulina*) and its role in the maintenance of contact with the mother. *Journal of Zoology*, *202* (4), 583-590. doi: 10.1111/j.1469-7998.1984.tb05055.x

Renouf, D. (1985). A demonstration of the ability of the harbour seal, *Phoca vitulina* (L.) to discriminate among pup vocalisations. *Journal of Experimental Marine Biology and Ecology*, *87* (1), 41-46. doi: 10.1016/0022-0981(85)90190-X

Revell, L. J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3* (2), 217-223. doi:10.1111/j.2041-210X.2011.00169.x

Reynolds, J. E., Wells, R. S., & Eide, S. D. (2000). *The Bottlenose Dolphin: Biology and Conservation* (pp. 260). University Press of Florida, US.

Richards, D. G., Wolz, J. P., & Herman, L. M. (1984). Vocal mimicry of computer-generated sounds and vocal labeling of objects by a bottlenosed dolphin, *Tursiops truncatus*. *Journal of Comparative Psychology*, *98* (1), 10. doi: 10.1037/0735-7036.98.1.10

Richardson, W.J, Greene, C.R., Malme, C.I., & Thomson, D.H. (1995). *Marine Mammals and Noise* (pp. 579). Academic Press, San Diego: California.

Ridgway, S. H., & Carder, D. A. (1993). High-frequency hearing loss in old (25+ years old) male dolphins. *The Journal of the Acoustical Society of America*, *94* (3), 1830-1830. doi: 10.1121/1.407769

Ridgway, S., Carder, D., Jeffries, M., & Todd, M. (2012). Spontaneous human speech mimicry by a cetacean. *Current Biology*, *22* (20), R860-R861. doi: 10.1016/j.cub.2012.08.044

Riebel, K. (2003). The "mute" sex revisited: vocal production and perception learning in female songbirds. *Advances in the Study of Behavior*, *33*, 49-86.

Ritschard, G., Bürgin, R. & Studer, M. (2013). Exploratory mining of life event histories. In *Contemporary Issues in Exploratory Data Mining in the Behavioral Sciences* (pp. 221-253). Routledge, US.

Robbins, R. L., & McCreery, E. (2003). African wild dog pup vocalisations with special reference to Morton's model. *Behaviour*, *140*, 333-351. doi: 10.1163/156853903321826666

Robinson, K. (2014). The role of oxytocin in the maternal behaviour of the grey seal (*Halichoerus grypus*) (pp. 238). Ph.D. thesis, University of St. Andrews.

Rogers, T. L., & Cato, D. H. (2002). Individual variation in the acoustic behaviour of the adult male leopard seal, *Hydrurga leptonyx*. *Behaviour*, *139* (10), 1267-1286. doi: 10.1163/156853902321104154

Rogers, T. L. (2003). Factors influencing the acoustic behaviour of male phocid seals. *Aquatic Mammals*, *29* (2), 247-260.

Rogers, T. L. (2007). Age-related differences in the acoustic characteristics of male leopard seals, *Hydrurga leptonyx*. *The Journal of the Acoustical Society of America*, *122* (1), 596-605. doi: 10.1121/1.2736976

Roitblat, H. L., Penner, R. H., & Nachtigall, P. E. (1990). Matching-to-sample by an echolocating dolphin (*Tursiops truncatus*). *Journal of Experimental Psychology: Animal Behavior Processes*, *16* (1), 85-95.doi: 10.1037/00977403.16.1.85

Romand, R., & Ehret, G. (1984). Development of sound production in normal, isolated, and deafened kittens during the first postnatal months. *Developmental Psychobiology*, *17* (6), 629-649. doi: 10.1002/dev.420170606

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology*, *1* (1), 43-46.

Roulin, A. (2003). The neuroendocrine function of allosuckling. *Ethology*, *109* (3), 185-195. doi: 10.1046/j.1439-0310.2003.00870.x

Roush, R.S., & Snowdon, C.T. (1994). Ontogeny of food-associated calls in cotton-top tamarins. *Animal Behaviour*, *47* (2), 263-273. doi: 10.1006/anbe.1994.1038

Roux, J. P., & Jouventin, P. (1987). Behavioral cues to individual recognition in the subantarctic fur seal, *Arctocephalus tropicalis*. *NOAA Tech. Rep. NMFS*, *51*, 95-102.

Rübsamen, R. (1987). Ontogenesis of the echolocation system in the rufous horseshoe bat, *Rhinolophus rouxi* (Audition and vocalisation in early postnatal development). *Journal of Comparative Physiology A*, *161* (6), 899-913. doi: 10.1007/BF00610231

Rübsamen, R., & Schäfer, M. (1990). Audiovocal interactions during development? Vocalisation in deafened young horseshoe bats vs. audition in vocalisation-impaired bats. *Journal of Comparative Physiology A, 167* (6), 771-784. doi: 10.1007/BF00189767

Ruddell, S. J. S., Twiss, S. D., & Pomeroy, P. P. (2007). Measuring opportunity for sociality: quantifying social stability in a colonially breeding phocid. *Animal Behaviour, 74* (5), 1357-1368. doi: 10.1016/j.anbehav.2007.01.024

Rukstalis, M., Fite, J. E., & French, J. A. (2003). Social change affects vocal structure in a callitrichid primate (*Callithrix kuhlii*). *Ethology, 109* (4), 327-340. doi: 10.1046/j.1439-0310.2003.00875.x

Sachs, M. (2014). cosinor: Tools for estimating and predicting the cosinor model. R package version 1.1. http://github.com/sachsmc/cosinor

Sales, G. D., & Smith, J. C. (1978). Comparative studies of the ultrasonic calls of infant murid rodents. *Developmental Psychobiology, 11* (6), 595-619. doi: 10.1002/dev.420110609

Salzinger, K., & Waller, M.B. (1962). The operant control of vocalisation in the dog. *Journal of the Experimental Analysis of Behavior, 5* (3), 383-389. doi: 10.1901/jeab.1962.5-383

Sanvito, S. & Galimberti, F. (2000). Bioacoustics of Southern elephant seals. II. Individual and geographical variation in male aggressive vocalisations. *Bioacoustics, 10*, 287-307. doi: 10.1080/09524622.2000.9753439

Sanvito, S., Galimberti, F., & Miller, E. H. (2007). Observational evidences of vocal learning in Southern elephant seals: a longitudinal study. *Ethology, 113* (2), 137-146. doi: 10.1111/j.1439-0310.2006.01306.x

Sanvito, S., Galimberti, F., & Miller, E. H. (2008). Development of aggressive vocalisations in male Southern elephant seals (*Mirounga leonina*): maturation or learning?. *Behaviour, 145* (2), 137-170. doi: 10.1163/156853907783244729

Savage-Rumbaugh, E. S. (1986). Ape language: from conditioned response to symbol (pp. 433). Columbia University Press, US.

Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., Rumbaugh, D. M., & Bates, E. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, i-252.

Sayigh, L. S., Tyack, P. L., Wells, R. S., & Scott, M. D. (1990). Signature whistles of free-ranging bottlenose dolphins, *Tursiops truncatus*: stability and mother-offspring comparisons. *Behavioral Ecology and Sociobiology*, *26* (4), 247-260. doi: 10.1007/BF00178318

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, *19* (10), 831-836. doi: 10.1016/j.cub.2009.03.061

Schevill, W. E., Watkins, W. A., & Ray, C. (1963). Underwater sounds of pinnipeds. *Science*, *141* (3575), 50-53. doi: 10.1126/science.141.3575.50

Schnitzler, H. U. (1973). Control of Doppler shift compensation in the greater horseshoe bat*, Rhinolophus ferrumequinum. Journal of Comparative Physiology*, *82* (1), 79-92. doi: 10.1007/BF00714171

Schrier, A. M., & Brady, P. M. (1987). Categorization of natural stimuli by monkeys (*Macaca mulatta*): Effects of stimulus set size and modification of exemplars. *Journal of Experimental Psychology: Animal Behavior Processes*, *13* (2), 136-143. doi: 10.1037/0097-7403.13.2.136

Schulte-Pelkum, N., Wieskotten, S., Hanke, W., Dehnhardt, G., & Mauck, B. (2007). Tracking of biogenic hydrodynamic trails in harbour seals (*Phoca vitulina*). *Journal of Experimental Biology*, *210* (5), 781-787. doi: 10.1242/jeb.02708

Schulz, T. M., & Bowen, W. D. (2004). Pinniped lactation strategies: evaluation of data on maternal and offspring life history traits. *Marine Mammal Science*, *20* (1), 86-114. doi: 10.1111/j.1748-7692.2004.tb01142.x

Schusterman, R. J. (1978). Vocal communication in pinnipeds. In: *Behavior of Captive Wild Animals* (pp. 247-308). Nelson-Hall, US.

Schusterman, R.J. (2008). Vocal learning in mammals with special emphasis on pinnipeds. In: *The Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication (*pp. 41-70). MIT Press, US.

Schusterman, R. J., & Feinstein, S. H. (1965). Shaping and discriminative control of underwater click vocalisations in a California sea lion. *Science*, *150* (3704), 1743-1744. doi: 10.1126/science.150.3704.1743

Schusterman, R. J., & Balliet, R. F. (1970). Conditioned vocalisations as a technique for determining visual acuity thresholds in sea lions. *Science*, *169* (3944), 498-501. doi: 10.1126/science.169.3944.498

Schusterman, R. J., Balliet, R. F., & John, S. S. (1970). Vocal displays under water by the gray seal, the harbor seal, and the stellar sea lion. *Psychonomic Science*, *18* (5), 303-305. doi: 10.3758/BF03331839

Schusterman, R.J., Balliet, R.F., & Nixon, J. (1972). Underwater audiogram of the California sea lion by the conditioned vocalisation technique. *Journal of Experimental Analysis of Behaviour*, *17* (3), 339-350. doi: 10.1901/jeab.1972.17-339

Schusterman, R.J., Hanggi, E., & Gisiner, R. (1992). Acoustic signalling in mother-pup reunions, interspecies bonding, and affiliation by kinship in California sea lions (*Zalophus californianus*). In: *Marine Mammal Sensory Systems* (pp. 533-551). Springer, US.

Schusterman, R. J., Kastak, D., Levenson, D. H., Reichmuth, C. J., & Southall, B. L. (2000). Why pinnipeds don't echolocate. *The Journal of the Acoustical Society of America*, *107* (4), 2256-2264. doi: 10.1121/1.428506

Schusterman, R. J., Reichmuth, C., & Kastak, D. (2002). The cognitive sea lions: Meaning and memory in the lab and in nature. In: *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* (pp. 217–228). MIT Press, US.

Schusterman, R. J., & Van Parijs, S. M. (2003). Pinniped vocal communication: an introduction. *Aquatic Mammals*, *29* (2), 177-180.

Schusterman, R. J., & Reichmuth, C. (2008). Novel sound production through contingency learning in the Pacific walrus (*Odobenus rosmarus divergens*). *Animal Cognition*, *11* (2), 319-327. doi: 10.1007/s10071-007-0120-5

Searcy, W.A., & Yasukawa, K. (1996). Song and female choice. In: *Ecology and Evolution of Acoustic Communication in Birds* (pp. 454-473). Cornell University Press, US.

Seyfarth, R.M., Cheney, D.L. & Marler, P. (1980a). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, *210* (4471), 801-803. doi:10.1126/science.7433999

Seyfarth, R.M., Cheney, D.L., & Marler, P. (1980b). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, *28*, 1070-1094. doi: 10.1016/50003-3472(80)80097-2

Seyfarth, R.M. & Cheney, D.L. (1986). Vocal development in vervet monkeys. *Animal Behaviour, 34* (6), 1640-1658. doi: 10.1016/S0003-3472(86)80252-4

Seyfarth, R.M., & Cheney, D.L. (1990). The assessment by vervet monkeys of their own and another species' alarm calls. *Animal Behaviour, 40* (4), 754-764. doi: 10.1016/S0003-3472(05)80704-3

Seyfarth, R. M., & Cheney, D. L. (1997). Some general features of vocal development in nonhuman primates. In: *Social Influences on Vocal Development.* Cambridge University Press, UK.

Seyfarth, R.M. & Cheney, D.L. (2010). Production, usage, and comprehension in animal vocalisations. *Brain and Language, 115* (1), 92-100. doi: 10.1016/j.bandl.2009.10.003

Shapiro, A.D., Slater, P.J.B., & Janik, V.M. (2004). Call usage learning in grey seals (*Halichoerus grypus*). *Journal of Comparative Psychology, 118* (4), 447-454. doi: 10.1037/0735-7036.118.4.447

Sherman, P.W. (1977). Nepotism and the evolution of alarm calls. *Science, 197*, 1246-1253. doi: 10.1126/science.197.4310.1246

Shipley, C., Hines, M., & Buchwald, J. S. (1981). Individual differences in threat calls of northern elephant seal bulls. *Animal Behaviour, 29* (1), 12-19. doi: 10.1016/S0003-3472(81)80147-9

Shipley, C., Buchwald, J. S., & Carterette, E. C. (1988). The role of auditory feedback in the vocalisations of cats. *Experimental Brain Research, 69* (2), 431-438. doi: 10.1007/BF00247589

Slabbekoorn, H., & den Boer-Visser, A. (2006). Cities change the songs of birds. *Current Biology, 16* (23), 2326-2331. doi: 10.1016/j.cub.2006.10.008

Slabbekoorn, H., Bouton, N., van Opzeeland, I., Coers, A., ten Cate, C., & Popper, A. N. (2010). A noisy spring: the impact of globally rising underwater sound levels on fish. *Trends in Ecology & Evolution, 25* (7), 419-427. doi: 10.1016/j.tree.2010.04.005

Slobodchikoff, C. N., & Coast, R. (1980). Dialects in the alarm calls of prairie dogs. *Behavioral Ecology and Sociobiology, 7* (1), 49-53. doi: 10.1007/BF00302518

Smiseth, P. T., & Lorentsen, S. H. (1995). Evidence of equal maternal investment in the sexes in the polygynous and sexually dimorphic grey seal (*Halichoerus grypus*). *Behavioral Ecology and Sociobiology, 36* (3), 145-150. doi: 10.1007/BF00177790

Smolker, R., & Pepper, J. W. (1999). Whistle convergence among allied male bottlenose dolphins (Delphinidae, Tursiops sp.). *Ethology, 105* (7), 595-617. doi: 10.1046/j.1439-0310.1999.00441.x

Snowdon, C. T. (1990). Language capacities of nonhuman animals. *American Journal of Physical Anthropology*, *33* (S11), 215-243. doi: 10.1002/ajpa.1330330510

Snowdon, C. T., & Hausberger, M. (Eds.). (1997). *Social Influences on Vocal Development* (pp. 363). Cambridge University Press, UK.

Snowdon, C.T., & Elowson, A.M. (1999). Pygmy marmosets modify call structure when paired. *Ethology*, *105*, 893-908. doi: 10.1046/j.1439-0310.1999.00483.x

Snowdon, C. T., & Elowson, A. M. (2001). 'Babbling' in pygmy marmosets: Development after infancy. *Behaviour*, *138* (10), 1235-1248. doi: 10.1163/15685390152822193

Soltis, J., Leong, K., & Savage, A. (2005). African elephant vocal communication I: antiphonal calling behaviour among affiliated females. *Animal Behaviour*, *70* (3), 579-587. doi: 10.1016/j.anbehav.2004.11.015

Somers, P. (1973). Dialects in southern Rocky Mountain pikas, *Ochotona princeps* (Lagomorpha). *Animal Behaviour*, *21* (1), 124-137. doi: 10.1016/S0003-3472(73)80050-8

Stanger, K. F., & Macedonia, J. M. (1994). Vocalisations of aye-ayes (*Daubentonia madagascariensis*) in captivity. *Folia Primatologica*, *62* (1-3), 160-169. doi: 0015-5713/94/0623-0160$8.00/0

Stansbury, A. L., Thomas, J. A., Stalf, C. E., Murphy, L. D., Lombardi, D., Carpenter, J., & Mueller, T. (2014). Behavioral audiogram of two Arctic foxes (*Vulpes lagopus*). *Polar Biology*, *37* (3), 417-422. doi: 10.1007/s00300-014-1446-5

Starr, R. M., O'Connell, V., Ralston, S., & Breaker, L. (2005). Use of acoustic tags to estimate natural mortality, spillover, and movements of lingcod (*Ophiodon elongatus*) in a marine reserve. *Marine Technology Society Journal*, *39* (1), 19-30. doi: 10.4031/002533205787521677

Sterbing, S. J. (2002). Postnatal development of vocalisations and hearing in the phyllostomid bat, *Carollia perspicillata*. *Journal of Mammalogy*, *83* (2), 516-525. doi: 10.1644/1545-1542(2002)083<0516:PDOVAH>2.0.CO;2

Sticken, J., & Dehnhardt, G. (2000). Salinity discrimination in harbour seals: a sensory basis for spatial orientation in the marine environment?. *Naturwissenschaften*, *87* (11), 499-502. doi: 10.1007/s001140050767

Stirling, I., & Warneke, R. M. (1971). Implications of a comparison of the airbourne vocalisations and some aspects of the behaviour of the two Australian fur seals, *Arctocephalus spp.*, on the evolution and present taxonomy of the genus. *Australian Journal of Zoology*, *19* (3), 227-241. doi: 10.1071/ZO9710227

Stirling, I. (1973). Vocalisation in the ringed seal (*Phoca hispida*). *Journal of the Fisheries Board of Canada*, *30* (10), 1592-1594. doi: 10.1139/f73-253

Stirling, I., Calvert, W., & Spencer, C. (1987). Evidence of stereotyped underwater vocalisations of male Atlantic walruses (*Odobenus rosmarus rosmarus*). *Canadian Journal of Zoology*, *65* (9), 2311-2321. doi: 10.1139/z87-348

Stoeger-Horwath, A. S., Stoeger, S., Schwammer, H. M., & Kratochvil, H. (2007). Call repertoire of infant African elephants: first insights into the early vocal ontogeny. *The Journal of the Acoustical Society of America*, *121* (6), 3922-3931. doi: 10.1121/1.2722216

Stoeger, A. S., Mietchen, D., Oh, S., de Silva, S., Herbst, C. T., Kwon, S., & Fitch, W. (2012). An Asian elephant imitates human speech. *Current Biology*, *22* (22), 2144-2148. doi: 10.1016/j.cub.2012.09.022

Suthers, R.A. (1999). The motor basis of vocal performance in songbirds. In: *The Design of Animal Communication* (pp. 37-62). MIT Press, US.

Suthers, R. A., & Fattu, J. M. (1973). Mechanisms of sound production by echolocating bats. *American Zoologist*, *13* (4), 1215-1226. doi: 10.1093/icb/13.4.1215

Suzuki, R., Buck, J. R., & Tyack, P. L. (1999). Information entropy of humpback whale song. *The Journal of the Acoustical Society of America*, *105* (2), 1048-1048. doi:10.1121/1.2161827

Symington, M. M. (1990). Fission-fusion social organization in Ateles and Pan. *International Journal of Primatology*, *11* (1), 47-61. doi: 10.1007/BF02193695

Taylor, A. M., Reby, D., & McComb, K. (2009). Context-related variation in the vocal growling behaviour of the domestic dog (*Canis familiaris*). *Ethology*, *115* (10), 905-915. doi: 10.1111/j.1439-0310.2009.01681.x

Taylor, A. M., & Reby, D. (2010). The contribution of source–filter theory to mammal vocal communication research. *Journal of Zoology*, *280* (3), 221-236. doi: 10.1111/j.1469-7998.2009.0061.x

Taylor, A. M., Reby, D., & McComb, K. (2010). Size communication in domestic dog, *Canis familiaris*, 'growls'. *Animal Behaviour, 79* (1), 205-210. doi: 101016/j.anbehav.2009.10.030

Tchernichovski, O., Nottebohm, F., Ho, C.E., Bijan, P., & Mitra, P.P. (2000). A procedure for an automated measurement of song similarity. *Animal Behaviour, 59* (6), 1167-1176. doi: 10.1006/anbe.1999.1416

Tembrock, G. (1976). Canid vocalisations. *Behavioural Processes, 1* (1), 57-75. doi: 10.1016/0376-6357(76)90007-3

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science, 308* (5730), 1934-1937. doi: 10.1126/science.1108841

ten Cate, C., & Bateson, P. (1988). Sexual selection: the evolution of conspicuous characteristics in birds by means of imprinting. *Evolution, 42* (6), 1355-1358.

Terhune, J. M., & Dell Apa, A. (2006). Stereotyped calling patterns of a male Weddell seal (*Leptonychotes weddellii*). *Aquatic Mammals, 32* (2), 175.

Terhune, J. M., Quin, D., Dell'Apa, A., Mirhaj, M., Plötz, J., Kindermann, L., & Bornemann, H. (2008). Geographic variations in underwater male Weddell seal trills suggest breeding area fidelity. *Polar Biology, 31* (6), 671-680. doi: 10.1007/s00300-008-0405-4

Terrace, H. S., Petitto, L. A., Sanders, R. J., & Bever, T. G. (1979). Can an ape create a sentence?. *Science, 206* (4421), 891-902. doi: 10.1126/science.504995

Thode, A., Straley, J., Tiemann, C. O., Folkert, K., & O'Connell, V. (2007). Observations of potential acoustic cues that attract sperm whales to longline fishing in the Gulf of Alaska. *The Journal of the Acoustical Society of America, 122* (2), 1265-1277. doi: 10.1121/1.2749450

Thomas, J. A., & Stirling, I. (1983). Geographic variation in the underwater vocalisations of Weddell seals (*Leptonychotes weddelli*) from Palmer Peninsula and McMurdo Sound, Antarctica. *Canadian Journal of Zoology, 61* (10), 2203-2212. doi: 10.1139/z83-291

Thomas, J. A., Zinnel, K. C., & Ferm, L. M. (1983). Analysis of Weddell seal (*Leptonychotes weddelli*) vocalisations using underwater playbacks. *Canadian Journal of Zoology, 61* (7), 1448-1456. doi: 10.1139/z83-194

Thorpe, W.H. (1964). Singing. In: *A New Dictionary of Birds* (pp. 739-750). Nelson, UK.

Titze, I.R. (1994). *Principles of Voice Production* (pp. 432). Prentice Hall, US.

Tokuda, I., Riede, T., Neubauer, J., Owren, M. J., & Herzel, H. (2002). Nonlinear analysis of irregular animal vocalisations. *The Journal of the Acoustical Society of America*, *111* (6), 2908-2919. doi: 10.1121/1.1474440

Trappe, M., & Schnitzler, H. U. (1982). Doppler-shift compensation in insect-catching horseshoe bats. *Naturwissenschaften*, *69* (4), 193-194. doi: 10.1007/BF00364902

Trillmich, F. (1981). Mutual mother-pup recognition in Galápagos fur seals and sea lions: Cues used and functional significance. *Behaviour, 78* (1-2), 21-42. doi: 10.1163/156853981X00248

Tripovich, J. S., Rogers, T. L., & Arnould, J. P. (2005). Species-specific characteristics and individual variation of the bark call produced by male Australian fur seals, *Arctocephalus pusillus doriferus*. *Bioacoustics*, *15* (1), 79-96. doi: 10.1080/09524622.2005.9753539

Tripovich, J. S., Charrier, I., Rogers, T. L., Canfield, R., & Arnould, J. P. (2008). Acoustic features involved in the neighbour–stranger vocal recognition process in male Australian fur seals. *Behavioural Processes*, *79* (1), 74-80. doi: 10.1016/j.beproc.2008.04.007

Tyack, P. (1986). Whistle repertoires of two bottlenosed dolphins, *Tursiops truncatus*: mimicry of signature whistles?. *Behavioral Ecology and Sociobiology*, *18* (4), 251-257. doi: 10.1007/BF00300001

Tyack, P.L. (1998). Acoustic communication under the sea. In: *Animal Acoustic Communication: Recent Technical Advances* (pp. 163-220). Springer-Verlag, EU.

Tyack, P. L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. *Journal of Comparative Psychology*, *122* (3), 319. doi: 10.1037/a0013087

Tyack, P. (2009). Acoustic playback experiments to study behavioral responses of free-ranging marine animals to anthropogenic sound. *Marine Ecology Progress Series, 395,* 187-200. doi: 10.3354/meps08363

Tyack, P. L., & Sayigh, L. S. (1997). Vocal learning in cetaceans. In: *Social Influences on Vocal Development* (pp. 208-233). Cambridge University Press, UK.

Tyack, P.L., & Janik, V.M. (2013). Effects of noise on acoustic signal production in marine mammals. In: *Effects of Noise on Animal Communication* (pp 251-271). Springer, US.

Tyson, R. B., Nowacek, D. P., & Miller, P. J. (2007). Nonlinear phenomena in the vocalisations of North Atlantic right whales (*Eubalaena glacialis*) and killer whales (*Orcinus orca*). *The Journal of the Acoustical Society of America*, *122* (3), 1365-1373. doi: 10.1121/1.2756263

Ulanovsky, N., Fenton, M. B., Tsoar, A., & Korine, C. (2004). Dynamics of jamming avoidance in echolocating bats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271* (1547), 1467-1475. doi: 10.1098/rspb.2004.2750

Van Opzeeland, I. C., & Van Parijs, S. M. (2004). Individuality in harp seal, *Phoca groenlandica*, pup vocalisations. *Animal Behaviour*, *68* (5), 1115-1123. doi: 10.1016/j.anbehav.2004.07.005

Van Opzeeland, I.C., Van Parijs, S.M., Frickenhaus, S., Kreiss, C.M., & Boebel, O. (2011). Individual variation in pup vocalisations and absence of behavioral signs of maternal vocal recognition in Weddell seals (*Leptonychotes weddellii*). *Marine Mammal Science*, E158-E172. doi: 10.1111/j.1748-7692.2011.00505.x

Van Parijs, S.M., Thompson, P.M., Tollit, D.J., & Mackay, A. (1997). Distribution and activity of male harbour seals during the mating season. *Animal Behaviour*, *54* (1), 35–43. doi: 10.1006/anbe.1996.0426

Van Parijs, S. M., & Corkeron, P. J. (2002). Ontogeny of vocalisations in infant black flying foxes, *Pteropus alecto*. *Behaviour*, *139* (9), 1111-1124. doi: 10.1163/15685390260437281

Vaňková, D., & Málek, J. (1997). Characteristics of the vocalisations of red deer Cervus elaphus hinds and calves. *Bioacoustics*, *7* (4), 281-289. doi: 10.1080/09524622.1997.9753340

Vater, M., Kössl, M., Foeller, E., Coro, F., Mora, E., & Russell, I. J. (2003). Development of echolocation calls in the mustached bat, *Pteronotus parnellii*. *Journal of Neurophysiology*, *90* (4), 2274-2290. doi: 10.1152/jn.00101.2003

Vergara, V. (2011). Acoustic communication and vocal learning in belugas (*Delphinapterus leucas*) (pp. 153). PhD thesis, The University of British Columbia.

Vergara, V., & Barrett-Lennard, L.G. (2008). Vocal development in a beluga calf (*Delphinapterus leucas*). *Aquatic Mammals*, *34* (1), 123-143. doi: 10.1578/AM.34.1.2008.123

Wargo-Rub, A.M., Gilbreath, L.G., McComas, R.L., Sandford, B., Teel, D.J., & Ferguson, J.W. (2012a). Estimated survival of adult spring/summer Chinnok salmon from the mouth of the Columbia river to Bonneville dam, 2011. *Report of the National Marine Fisheries Service*, Northwest Fisheries Science Center, Seattle, pp. 1-28.

Wargo-Rub, A.M., Sandford, B.P., Gilbreath, L.G., Myers, M.S., Peterson, M.E., Charlton, L.L., Smith, S.G., Matthews, G.M. (2012b). Comparative performance of acoustic tagged and passive integrated transponder tagged juvenile Chinook salmon in the Columbia and Snake rivers, 2008. *Report of the National Marine Fisheries Service to the U.S. Army Corps of Engineers*, Portland District, Portland Oregon, pp 1-178.

Wasserman, E.A., Zentall, T.R., & Cook, R.G. (2006). Relational discrimination learning in pigeons. In: *Comparative Cognition* (pp. 307-324).Oxford University Press, US.

Watanabe, S. (2001). Van Gogh, Chagall and pigeons: picture discrimination in pigeons and humans. *Animal Cognition*, *4* (3-4), 147-151. doi: 10.1007/s100710100112

Watkins, W.A., Moore, K.E., Clark, C.W., & Dahlheim, M.E. (1988). The sounds of sperm whale calves. In: *Animal Sonar* (pp.99-107). Springer, US.

Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., West, V. & Slocombe, K. E. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology, 25*, 495-499. doi: 10.1016/j.cub.2014.12.032

Watwood, S. L., Tyack, P. L., & Wells, R. S. (2004). Whistle sharing in paired male bottlenose dolphins, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, *55* (6), 531-543. doi: 10.1007/s00265-003-0724-y

Weary, D. M., & Krebs, J. R. (1992). Great tits classify songs by individual voice characteristics. *Animal Behaviour, 43* (2), 283-287. doi: 10.1016/S0003-3472(05)80223-4

Weckerly, F. (1998). Sexual-size dimorphism: Influence of mass and mating systems in the most dimorphic mammals. *Journal of Mammalogy*, *79* (1), 33-52. doi:10.2307/1382840

Weedon, J.S., & Falls, J.B. (1959). Differential response of male ovenbirds to recorded songs of neighbouring and more distant individuals. *The Auk, 76*, 343-351.

Weiss, D. J., Garibaldi, B. T., & Hauser, M. D. (2001). The production and perception of long calls by cotton-top tamarins (*Saguinus oedipus*): acoustic analyses and playback experiments. *Journal of Comparative Psychology*, *115* (3), 258. doi: 10.1037/0735-7036.115.3.258

Wells, R. S. (1991). The role of long-term study in understanding the social structure of a bottlenose dolphin community. In: *Dolphin Societies: Discoveries and Puzzles*, (pp. 199-225). University of California Press, US.

Wich, S. A., Swartz, K. B., Hardus, M. E., Lameira, A. R., Stromberg, E., & Shumaker, R. W. (2009). A case of spontaneous acquisition of a human sound by an orangutan. *Primates*, *50* (1), 56-64. doi: 10.1007/s10329-008-0117-y

Wickler, W. (2013). Understanding mimicry–with special reference to vocal mimicry. *Ethology*, *119* (4), 259-269. doi: 10.1111/eth.12061

Williams, S. L., Brakke, K. E., & Savage-Rumbaugh, E. S. (1997). Comprehension skills of language-competent and nonlanguage-competent apes. *Language & Communication*, *17* (4), 301-317. doi: 10.1016/S0271-5309(97)00012-8

Wilson, W. A. (1975). Discriminative conditioning of vocalisations in *Lemur catta*. *Animal Behaviour*, *23*, 432-436. doi: 10.1016/0003-3472(75)90091-3

Winter, P., Handley, P., Ploog, D., & Schott, D. (1973). Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour*, *47* (3), 230. doi: 10.1163/156853973X00085

Wolski, L. F., Anderson, R. C., Bowles, A. E., & Yochem, P. K. (2003). Measuring hearing in the harbor seal (*Phoca vitulina*): Comparison of behavioral and auditory brainstem response techniques. *The Journal of the Acoustical Society of America*, *113* (1), 629-637. doi: 10.1121/1.1527961

Woolf, N. (1974). The ontogeny of bat sonar sounds: with special emphasis on sensory deprivation (pp. 140). Sc.D thesis, Johns Hopkins University.

Wyman, M. T., Mooring, M. S., McCowan, B., Penedo, M. C. T., Reby, D., & Hart, L. A. (2012). Acoustic cues to size and quality in the vocalisations of male North American bison, *Bison bison*. *Animal Behaviour*, *84* (6), 1381-1391. doi: 10.1016/j.anbehav.2012.08.037

Zann, R. (1994). Effects of band color on survivorship, body condition and reproductive effort on free-living Australian zebra finches. *The Auk, 111*, 131-142. doi: 82.43.98.187

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, *102* (6), 3191-3202. doi: 10.1152/jn.00066.2009

Zhang, L., Jones, G., Parsons, S., Liang, B., & Zhang, S. (2005). Development of vocalisations in the flat-headed bats, *Tylonycteris pachypus* and *T. robustula* (Chiropter: Vespertilionidae). *Acta Chiropterologica*, *7*(1), 91-99.

Zuberbühler, K. (2001). Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus campbelli*. *Behavioral Ecology and Sociobiology*, *50* (5), 414-422. doi: 10.1007/s002650100383

Zuur, A.F., Leno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M. (2009). *Mixed Effects Models and Extensions in Ecology with R* (pp. 574). Springer, US.

Appendix

Appendix I. Example recordings of grey seal pup (A through D), adult female (F through H) and adult male (I and J) calls, corresponding to the spectrograms shown in Figure 3.2. Audio files are attached in the supplemental material CD.

Appendix II. Recording of underwater call sequence shown in Figure 3.3a. The audio file is attached in the supplemental material CD.

Appendix III. Mean and standard deviation for the average fundamental frequency, peak frequency and duration by call type, age class, and sex. Age was categorized into three categories; nursing pups 'Pup' (stages 1-4), weaned pups 'Weaned' (stage 5 to 1 year) and Adults (older than 1 year).

Noisy			
Peak (Hz)	**Pup**	**Weaned**	**Adult**
Male	551.220 (± 281.401)	528.931 (± 176.398)	431.346 (± 193.863)
Female	609.981 (± 196.657)	574.849 (± 162.05)	486.609 (± 162.518)
Duration (Sec)			
Male	0.371 (± 0.269)	0.664 (± 0.548)	1.646 (± 1.719)
Female	0.510 (± 0.422)	0.711 (± 0.670)	1.236 (± 1.054)

Periodic			
Fundamental (Hz)	**Pup**	**Weaned**	**Adult**
Male	533.344 (±96.831)	401.674 (± 114.811)	272.947 (± 70.665)
Female	602.146 (± 99.702)	487.628 (± 125.477)	368.366 (± 99.649)
Peak (Hz)			
Male	619.136 (±211.926)	408.179 (± 203.271)	274.211 (± 71.085)
Female	697.649 (± 187.864)	614.286 (± 300.856)	362 (± 84.774)
Duration (Sec)			
Male	0.805 (± 0.342)	1.269 (± 0.604)	2.308 (± 0.993)
Female	0.896 (± 0.375)	1.401 (± 0.629)	1.579 (± 1.291)

Mixed			
Peak (Hz)	**Pup**	**Weaned**	**Adult**
Male	661.861 (± 247.626)	428.391 (± 180.576)	353 (± 100.661)
Female	735.569 (± 242.162)	863 (NA)	417.736 (± 177.966)
Duration (Sec)			
Male	1.026 (± 0.507)	0.89 (± 0.423)	2.303 (± 1.826)
Female	1.026 (± 0.391)	0.928 (NA)	2.015 (± 1.355)

Appendix IV. Phylogenetic PCA loadings for 'noisy' calls.

Noisy	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
% Nonlinear	-0.19945	-0.25718	0.111293	-0.22453	-0.1992	0.202416	0.018461	-0.06783	-0.08289	-0.07567
# Peaks Max	-0.78589	-0.15089	0.044713	0.024308	0.125707	0.107719	0.091612	-0.05752	0.025449	0.02783
# Peaks Mean	-0.74353	-0.43111	0.068165	0.04099	0.152007	0.006276	0.011527	-0.05377	-0.08266	-0.01105
# Peaks Med	-0.37101	-0.07949	-0.11912	0.047313	0.235402	-0.12983	-0.03451	-0.161	0.202863	0.078885
# Peaks Q0	-0.48653	-0.07297	-0.24099	0.213943	0.073835	0.118981	0.091878	0.35766	-0.15869	0.191043
# Peaks Q1	-0.77541	-0.23801	0.017259	0.09077	-0.02982	-0.03577	0.072517	-0.05456	-0.04502	-0.0335
# Peaks Q2	-0.76833	-0.30899	0.078242	0.083686	0.069541	-0.08743	0.018233	-0.14167	-0.06433	-0.00637
# Peaks Q3	-0.73756	-0.26362	0.117132	0.011473	0.179146	-0.05339	-0.00509	-0.04578	0.004354	-0.05277
# Peaks Q4	-0.50125	-0.04106	-0.12317	0.037591	0.351126	0.081692	-0.18109	0.386966	-0.02508	-0.1157
1st Freq Peak Max	0.05761	0.329909	-0.23971	-0.24378	0.162006	-0.20641	0.013701	-0.10293	-0.17995	-0.16133
1st Freq Peak Mean	0.171333	0.415925	-0.41501	-0.26129	0.337346	-0.29048	0.135987	-0.23816	-0.2632	-0.16936
1st Freq Peak Med	-0.16647	0.325557	-0.07827	-0.26712	0.059242	0.126183	0.171263	0.015702	-0.08529	-0.1583
1st Freq Peak Q0	0.065481	0.198041	-0.26372	-0.18912	0.191057	-0.14875	0.160376	-0.17096	-0.13529	-0.07159
1st Freq Peak Q2	0.126017	0.279751	-0.33224	-0.25602	0.257541	-0.1793	0.088355	-0.20203	-0.17479	-0.12075
1st Freq Peak Q3	0.127193	0.316361	-0.29378	-0.13019	0.297492	-0.29079	0.076902	-0.15976	-0.22082	-0.15723
1st Freq Peak Q4	0.065507	0.159064	-0.22748	-0.13118	0.248449	-0.09373	0.138748	-0.32105	-0.25168	-0.11433
1st Freq Peak SD	-0.15801	0.061973	0.026854	-0.19613	-0.16994	0.128137	-0.05082	0.103942	0.083392	-0.1505
BW Max	-0.37236	0.050814	0.178476	-0.16396	-0.07751	0.386506	0.24306	-0.17881	-0.1867	-0.14714
BW Mean	-0.47001	-0.03919	0.09234	-0.18086	0.007175	0.423729	0.335708	-0.11445	-0.01836	-0.1553
BW Med	-0.73179	0.11957	0.19352	0.009498	-0.09682	0.171638	-0.01867	0.043865	-0.049	-0.07127
BW Q0	-0.1503	-0.02185	-0.25173	0.083676	0.021965	0.244718	0.118321	0.080333	-0.24425	0.245798
BW Q1	-0.38435	-0.06252	-0.04008	0.011298	-0.27078	0.237264	0.172671	0.014445	-0.08056	-0.33859
BW Q2	-0.34039	-0.02304	0.168602	-0.1652	-0.0687	0.251511	0.050758	-0.14893	-0.28889	-0.09169
BW Q3	-0.40962	0.030669	0.155327	-0.11306	-0.01142	0.164394	0.033776	-0.06011	0.017264	-0.18059
BW Q4	-0.2398	-0.01644	-0.12847	-0.06109	0.170095	0.308568	-0.11028	0.189443	-0.121	-0.10891
BW SD	-0.46607	-0.06492	0.113661	-0.09651	-0.23787	0.294977	-0.01919	-0.15855	-0.22234	-0.08704
Dist to Max	0.095702	0.058192	-0.01968	-0.10245	-0.30372	-0.04899	-0.39183	0.292036	-0.53268	-0.01308

Noisy	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Duration	0.15597	0.058168	0.009832	-0.14343	-0.36344	-0.03857	-0.41012	0.271163	-0.55359	-0.04488
Energy	0.116997	0.072059	-0.13525	-0.32251	-0.27868	-0.18224	-0.24919	0.135816	-0.39074	-0.03805
Entropy Max	-0.83605	-0.26131	0.074532	-0.02055	0.128889	0.096681	0.141869	-0.09474	-0.02881	-0.01127
Entropy Mean	-0.89932	-0.34347	0.036549	0.056964	0.08492	-0.04065	0.018709	-0.00874	-0.06361	-0.02426
Entropy Med	-0.89778	-0.31988	0.032661	0.016596	0.055878	-0.06452	-0.01013	-0.0869	-0.0576	-0.0131
Entropy Q0	-0.56099	-0.09756	-0.26835	0.178724	0.074391	0.104647	0.146236	0.364954	-0.20388	0.206258
Entropy Q1	-0.82035	-0.29498	0.014059	0.078594	-0.04051	-0.03302	0.061848	-0.04472	-0.05473	-0.02734
Entropy Q2	-0.80512	-0.36401	0.095269	0.036658	0.037819	-0.08596	0.008782	-0.14424	-0.10185	-0.02156
Entropy Q3	-0.77867	-0.3225	0.126576	-0.00833	0.142311	-0.0535	-0.01389	-0.0521	0.017956	-0.05744
Entropy Q4	-0.54426	-0.07039	-0.1118	-0.00713	0.353676	0.058229	-0.15217	0.382636	-0.06802	-0.10049
Entropy SD	0.403	0.056553	-0.13645	-0.01137	0.073026	-0.4681	-0.23594	-0.09116	-0.04985	0.089911
HNR Max	0.423872	0.234219	0.070922	0.346669	-0.05486	0.030361	-0.13048	0.037576	-0.01806	-0.03599
HNR Mean	0.521621	0.359786	0.155593	0.594859	0.1272	0.081941	-0.03094	-0.13756	-0.21505	-0.13189
HNR Med	0.270727	0.305269	0.159438	0.576417	0.131973	0.091289	0.071131	-0.07825	-0.19598	-0.12837
HNR Q0	0.225234	0.117362	0.259498	0.301193	0.128005	-0.07615	0.008771	-0.28031	-0.12626	-0.25413
HNR Q1	0.39737	0.254392	0.148648	0.380266	0.154839	0.05344	-0.06479	-0.09185	-0.16752	-0.08773
HNR Q2	0.378737	0.30704	0.089387	0.423077	0.108508	0.063675	-0.01507	-0.01267	-0.15378	-0.06283
HNR Q3	0.331641	0.29539	0.062188	0.472807	0.019518	0.123094	-0.0146	-0.05318	-0.19086	-0.0183
HNR Q4	0.185253	0.144451	0.164971	0.383865	0.001657	0.129548	0.053016	-0.27527	-0.1348	-0.04418
HNR SD	-0.31868	-0.28862	-0.12026	-0.5087	-0.28106	-0.04203	-0.03295	-0.05185	0.178481	0.032588
Max Freq Max	-0.5104	0.544156	0.269531	-0.20478	0.022718	0.176001	0.079599	-0.10981	-0.15658	0.212125
Max Freq Mean	-0.73608	0.607329	-0.07943	0.044933	-0.1401	-0.01382	-0.02023	-0.00731	0.080157	-0.0358
Max Freq Med	-0.74973	0.214003	0.195136	0.012375	-0.08516	0.109069	-0.01905	0.03744	-0.035	-0.05229
Max Freq Q0	-0.20833	0.077262	-0.71839	0.206277	-0.1923	0.195196	0.186713	-0.13797	-0.08415	0.401358
Max Freq Q1	-0.51374	0.248891	-0.3615	0.213569	-0.40357	-0.08556	0.146208	0.105617	0.132227	-0.46923
Max Freq Q2	-0.57551	0.599165	0.258073	-0.06416	-0.06802	-0.0416	0.028676	-0.06329	-0.04335	0.088623
Max Freq Q3	-0.5503	0.50155	0.168188	-0.0243	0.012107	-0.06167	-0.19676	0.079028	0.151732	0.049454
Max Freq Q4	-0.30684	0.186192	-0.41135	-0.09564	0.41776	0.433028	-0.44582	-0.02858	0.111836	-0.14096

257

Noisy	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Max Freq SD	-0.46642	-0.28988	-0.03547	0.260395	-0.27572	-0.02975	-0.4896	-0.32965	-0.03923	0.066131
Min Freq Max	-0.44881	0.587111	0.242686	-0.17552	0.04945	0.073817	0.01194	-0.06559	-0.11494	0.281212
Min Freq Mean	-0.68879	0.602425	-0.08185	0.090586	-0.17867	-0.09307	-0.12154	-0.02343	0.095253	-0.00651
Min Freq Med	-0.19538	0.571785	0.033185	0.017935	0.056905	-0.34814	-0.00441	-0.03249	0.076612	0.103366
Min Freq Q0	-0.16798	0.097212	-0.69901	0.195602	-0.22731	0.110619	0.157976	-0.19211	0.014633	0.343272
Min Freq Q1	-0.44816	0.298422	-0.39061	0.234781	-0.36224	-0.17304	0.106877	0.113255	0.173982	-0.41345
Min Freq Q2	-0.50323	0.660529	0.219904	-0.01005	-0.04917	-0.13633	0.012869	-0.01498	0.057402	0.12964
Min Freq Q3	-0.47626	0.547891	0.135577	0.010444	0.017281	-0.12319	-0.23012	0.107859	0.16305	0.114976
Min Freq Q4	-0.22803	0.217837	-0.40037	-0.07779	0.387182	0.336547	-0.44803	-0.12512	0.185307	-0.10536
Min Freq SD	-0.49074	-0.32563	-0.07689	0.239541	-0.20128	0.010339	-0.43475	-0.30472	-0.07028	0.046063
Peak Freq Max	-0.47879	0.574024	0.25178	-0.19909	0.040006	0.127196	0.045959	-0.08895	-0.13639	0.250372
Peak Freq Mean	-0.71477	0.62845	-0.09261	0.056743	-0.14456	-0.04785	-0.04912	0.001485	0.078207	-0.02819
Peak Freq Med	-0.49027	0.676421	0.178098	-0.09657	-0.06249	-0.03405	0.093441	0.047553	0.049391	0.083096
Peak Freq Q0	-0.19011	0.092977	-0.72179	0.186598	-0.21133	0.144538	0.18052	-0.17191	-0.01376	0.375233
Peak Freq Q1	-0.47075	0.283703	-0.38713	0.223254	-0.3809	-0.13374	0.12628	0.113551	0.163426	-0.44707
Peak Freq Q2	-0.53637	0.649094	0.239164	-0.03064	-0.05481	-0.08046	0.020545	-0.03077	0.013869	0.11281
Peak Freq Q3	-0.51026	0.541871	0.150188	-0.00698	0.012099	-0.09214	-0.21788	0.095444	0.159831	0.077507
Peak Freq Q4	-0.25858	0.217425	-0.40689	-0.09891	0.402858	0.37867	-0.45657	-0.10415	0.159536	-0.1213
Peak Freq SD	-0.40437	-0.31909	-0.01319	0.26877	-0.2692	-0.02872	-0.49517	-0.32827	-0.0086	0.083059
Peak to Peak	-0.16275	-0.10458	-0.2502	-0.45801	-0.25753	-0.2263	-0.13815	-0.11378	-0.2353	-0.0317
Total BW Max	-0.71499	-0.12856	0.036306	0.065089	0.169364	-0.12631	0.085318	-0.04863	-0.05252	0.027231
Total BW Mean	-0.83222	-0.13136	-0.04961	0.141789	0.098681	-0.19161	0.035344	0.007332	0.012823	0.01695
Total BW Medium	-0.85106	-0.20251	0.005037	0.095474	0.084161	-0.24575	0.021023	-0.05126	-0.04145	-0.01795
Total BW Q0	-0.42114	-0.05812	-0.19231	0.235237	0.112212	-0.13536	0.186888	0.395154	-0.14953	0.109828
Total BW Q1	-0.62429	-0.25533	0.013426	0.148088	0.031519	-0.28654	0.050409	0.048077	-0.09851	-0.07038
Total BW Q2	-0.61915	-0.24121	0.053213	0.166006	0.146806	-0.29602	0.035787	0.019715	-0.12323	0.04644
Total BW Q3	-0.59014	-0.22143	0.040024	0.123963	0.165909	-0.28411	0.028366	0.037103	-0.04856	-0.00368
Total BW Q4	-0.31002	-0.01669	-0.03196	0.113963	0.340972	-0.08162	-0.01826	0.356752	-0.13053	0.043051
Total BW SD	-0.12346	-0.31076	0.085418	-0.12093	0.053431	-0.29841	-0.07343	-0.21906	0.041527	0.033792

Appendix V. Phylogenetic PCA loadings for 'periodic' calls.

Periodic	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
% Nonlinear	0.44395	0.068405	0.074323	0.048558	0.077498	0.156188	-0.22869	-0.24208	0.040865	-0.12071
# Peaks Max	-0.23163	-0.22344	0.072332	-0.07414	-0.01377	0.086055	-0.01361	-0.10777	0.030324	0.340244
# Peaks Mean	0.075505	0.345406	-0.24042	-0.15948	-0.17647	0.155697	0.047137	0.037664	-0.302	-0.08135
# Peaks Med	-0.51902	-0.53086	0.015899	0.001155	0.008961	0.015737	0.001792	-0.0747	0.124988	0.272353
# Peaks Q0	0.610258	0.077654	0.076344	0.267201	0.10845	0.003569	-0.05058	-0.03063	0.001578	0.048714
# Peaks Q1	0.497187	-0.45104	-0.14326	-0.14506	-0.05047	-0.10271	0.209909	-0.1685	0.03195	-0.0707
# Peaks Q2	0.472826	0.113353	0.030601	0.100883	0.119562	-0.23682	-0.14446	-0.04038	-0.20488	-0.016
# Peaks Q3	0.52831	-0.40066	-0.1541	-0.13115	-0.05565	-0.11959	0.208766	-0.14542	0.028926	-0.12505
# Peaks Q4	0.434267	-0.49155	-0.12471	-0.16125	-0.06129	-0.08463	0.203282	-0.18916	0.041234	0.013274
1st Freq Peak Max	0.744469	-0.01343	-0.16547	0.11904	-0.03621	-0.0179	0.089468	0.064447	0.15648	-0.08341
1st Freq Peak Mean	-0.36509	-0.59773	0.177415	-0.21518	0.166097	0.083558	-0.03225	-0.16948	-0.2628	0.037059
1st Freq Peak Med	-0.02857	-0.69971	0.117896	-0.18625	0.173292	0.087338	0.01003	-0.16221	-0.22161	-0.00127
1st Freq Peak Q0	-0.19356	-0.45648	-0.07699	-0.194	-0.00943	0.050906	-0.04051	-0.11595	0.107162	-0.15346
1st Freq Peak Q2	0.499319	-0.06751	0.049055	0.21646	0.167033	-0.11751	-0.04765	0.027409	-0.1534	0.161859
1st Freq Peak Q3	0.639314	-0.4533	-0.09232	-0.03236	0.127634	0.029402	0.139346	-0.01635	-0.02962	0.022424
1st Freq Peak Q4	0.318702	-0.0844	0.096633	0.157061	0.294368	-0.20026	-0.12957	-0.07385	-0.36064	0.321509
1st Freq Peak SD	-0.50564	-0.74762	0.073133	-0.03942	0.088224	-0.05835	-0.07477	-0.04291	-0.06398	-0.00925
BW Max	-0.34356	-0.0112	-0.66887	0.301001	0.139057	0.00663	-0.09374	0.005769	-0.0181	-0.00651
BW Mean	-0.30572	-0.12812	-0.52964	0.134296	0.146282	-0.05328	-0.22982	-0.10072	0.034146	-0.06632
BW Med	-0.35286	-0.03113	-0.65197	0.18896	0.176271	-0.02817	-0.14555	-0.06643	0.015619	-0.04266
BW Q0	0.190228	-0.09812	0.105866	-0.20505	0.011681	0.033785	-0.21003	-0.16243	-0.06459	-0.13438
BW Q1	-0.16175	-0.16794	-0.35973	0.075654	0.10408	-0.07111	-0.22191	-0.08731	0.092714	-0.1106
BW Q2	-0.33056	-0.04804	-0.72731	0.197919	0.16704	-0.03348	-0.193	-0.08013	0.048555	-0.01617
BW Q3	-0.20188	-0.25922	-0.4127	0.113936	0.081079	-0.01807	-0.22913	-0.08449	0.07148	-0.10674
BW Q4	-0.26294	-0.13675	-0.14181	0.17892	0.04973	-0.17903	-0.17566	-0.03106	0.057822	0.078771
BW SD	-0.32427	-0.03897	-0.67597	0.229821	0.144993	-0.0444	-0.14959	-0.08139	0.101102	0.065696
Dist to Max	-0.34926	-0.01745	-0.69224	0.287917	0.160373	-0.02311	-0.08296	-0.03668	0.023714	0.051833
Duration	0.03665	0.005359	0.044335	-0.10677	-0.12861	-0.45565	-0.16914	0.083821	-0.3593	-0.39638

Periodic	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Energy	0.015253	0.013825	0.055297	-0.08053	-0.08756	-0.41017	-0.15285	0.051567	-0.38493	-0.44304
Entropy Max	0.554458	-0.56022	-0.17282	-0.09523	-0.01238	-0.02071	0.169713	0.025321	0.0063	0.0034
Entropy Mean	-0.53716	-0.5709	0.011322	-0.05952	0.147853	0.047954	-0.09728	-0.11106	0.067042	-0.02169
Entropy Med	-0.24093	-0.14915	0.123359	-0.24482	0.03977	0.093507	-0.12481	-0.17676	0.013303	0.122924
Entropy Q0	0.281529	0.064027	0.00675	-0.01453	-0.03723	-0.35687	-0.21751	0.002517	-0.10083	-0.36109
Entropy Q1	0.425389	-0.03873	0.106119	-0.07024	-0.04821	-0.31055	-0.28867	-0.01515	0.059898	-0.26796
Entropy Q2	0.576222	-0.54083	-0.18333	-0.0647	-0.01261	-0.02759	0.182443	0.042353	0.003601	-0.01046
Entropy Q3	0.401466	-0.05393	0.030203	0.193591	0.214208	-0.33461	-0.11596	0.078826	-0.27178	0.338125
Entropy Q4	0.587542	-0.5231	-0.19181	-0.04943	-0.01926	-0.03714	0.189094	0.056668	0.003888	-0.01878
Entropy SD	0.183894	0.338712	-0.14331	-0.0613	-0.17608	0.166412	0.05973	-0.05239	-0.20703	-0.01966
Fundamental Max	0.68531	-0.63169	-0.13195	-0.0826	-0.13888	0.012211	0.017781	0.040364	0.008174	-0.0064
Fundamental Mean	-0.19523	-0.2987	0.073096	-0.20682	0.148672	0.323132	-0.11462	-0.22866	0.053827	0.2084
Fundamental Med	0.681402	-0.65124	-0.13063	-0.07859	-0.08905	0.063597	0.005499	-0.00088	0.01209	-0.00447
Fundamental Q0	-0.07811	0.311402	-0.41768	-0.19046	-0.12682	0.256933	0.098748	-0.00183	-0.33887	0.044363
Fundamental Q1	0.185392	-0.23539	-0.23187	-0.1759	-0.26632	-0.14631	0.008701	0.12792	0.301917	-0.16743
Fundamental Q2	0.425715	-0.22429	0.200166	0.187276	0.301543	0.121131	-0.0601	-0.20281	-0.25134	0.097108
Fundamental Q3	0.693308	-0.64226	-0.13077	-0.06574	-0.0968	0.048288	0.013355	0.011649	0.006468	-0.01251
Fundamental Q4	0.491926	-0.21571	0.088442	0.131281	0.228299	-0.32715	-0.18971	0.000394	-0.28245	0.316337
Fundamental SD	-0.65592	-0.66721	0.03212	0.089342	0.081606	0.000203	0.020266	-0.00702	-0.05896	0.042517
HNR Max	-0.45529	-0.33742	-0.02673	-0.03638	-0.54377	-0.21296	-0.22722	0.13175	0.001684	0.153736
HNR Mean	-0.29764	-0.33296	0.190652	-0.26515	0.04583	0.070962	-0.13804	-0.15787	-0.11722	0.197185
HNR Med	-0.44415	-0.32722	-0.02087	-0.07914	-0.57873	-0.27447	-0.18584	0.186474	-0.00535	0.170556
HNR Q0	0.15176	0.571674	-0.38525	-0.2735	-0.21793	0.22967	0.064012	-0.00902	-0.3982	0.032344
HNR Q1	-0.49058	-0.722	0.069661	0.010341	-0.0325	-0.01491	0.05357	0.008677	-0.07674	-0.05607
HNR Q2	0.803741	0.073605	0.094306	0.239293	0.119086	0.069029	-0.14166	-0.14834	-0.00312	-0.07431
HNR Q3	-0.47875	-0.25983	-0.01766	-0.10354	-0.5848	-0.27615	-0.19887	0.183317	0.013743	0.172227
HNR Q4	-0.16354	-0.22442	-0.07923	-0.32641	-0.1886	-0.3497	-0.17791	0.133104	-0.11028	0.304111
HNR SD	0.207766	0.213964	0.036909	0.434189	-0.15375	-0.14184	0.155851	0.229839	-0.1268	0.049768
Max Freq Max	0.253859	-0.34638	0.116048	0.243559	-0.09539	0.468123	-0.4208	0.432115	-0.08449	-0.05016

Periodic	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Max Freq Mean	-0.37381	-0.38852	0.033952	0.427968	0.122991	0.127148	0.13568	0.292069	-0.20571	-0.136
Max Freq Med	-0.12655	-0.25874	0.095193	0.252117	0.065545	0.232047	0.026747	0.221352	-0.17747	-0.12782
Max Freq Q0	0.080641	-0.36149	0.357019	0.17694	0.068148	-0.19686	-0.09255	-0.02423	0.271328	-0.09676
Max Freq Q1	-0.2427	-0.21806	0.06708	-0.05834	0.040436	-0.19504	0.188948	0.200624	-0.15125	0.209273
Max Freq Q2	0.352217	-0.28875	0.09532	0.169723	-0.12565	0.435625	-0.48007	0.401561	-0.01808	-0.00187
Max Freq Q3	0.336716	0.142538	-0.00395	-0.15441	-0.01241	-0.07221	-0.3295	-0.15663	0.099069	0.099923
Max Freq Q4	0.37659	-0.22519	0.069809	0.111328	-0.15357	0.383034	-0.5018	0.348781	0.021602	0.027795
Max Freq SD	0.061443	0.292185	-0.24031	-0.26396	-0.10039	0.101514	-0.03404	-0.0784	-0.31621	0.09389
Min Freq Max	0.457846	-0.55137	-0.08947	-0.17178	0.018875	0.155369	-0.04148	0.214987	0.032761	0.049776
Min Freq Mean	-0.52495	-0.51038	0.030138	-0.01269	0.132782	0.025366	-0.0219	0.012719	-0.1783	-0.15303
Min Freq Med	-0.18527	-0.28557	0.118089	-0.15617	0.105877	0.142447	-0.03286	-0.10824	-0.22423	-0.0657
Min Freq Q0	-0.28269	-0.38482	0.089672	0.405451	0.066693	0.150306	0.145079	0.305428	-0.21568	-0.15114
Min Freq Q1	0.40782	0.058169	0.076629	0.022485	0.065895	0.123589	-0.23246	-0.24739	0.134222	-0.06943
Min Freq Q2	0.522427	-0.50143	-0.12221	-0.1369	-0.00415	0.128963	-0.03036	0.246859	0.082922	0.069611
Min Freq Q3	0.44564	0.144338	-0.00624	0.112421	0.056787	-0.18325	-0.12011	0.009127	0.032129	0.325832
Min Freq Q4	0.554175	-0.43967	-0.14493	-0.10696	-0.02799	0.096079	-0.02799	0.269738	0.134124	0.080588
Min Freq SD	0.091962	0.347641	-0.27468	-0.19331	-0.16661	0.149063	0.05673	-0.01704	-0.28486	0.07441
Peak Freq Max	-0.16941	-0.1944	0.150102	-0.2085	0.12205	0.094471	-0.11287	-0.17544	-0.17778	-0.05792
Peak Freq Mean	0.144858	0.345097	-0.262	-0.12423	-0.21498	0.167423	0.102485	0.042519	-0.21903	0.058106
Peak Freq Med	-0.55491	-0.55911	0.067107	-0.11311	0.115698	0.00564	-0.12564	-0.10294	-0.0793	-0.10278
Peak Freq Q0	-0.35025	-0.50903	0.034448	-0.07592	0.057169	0.011838	-0.02916	0.058906	-0.19629	-0.24198
Peak Freq Q1	0.514715	-0.52843	-0.24995	-0.16787	0.019124	-0.04216	0.148034	0.016051	-0.0147	0.011609
Peak Freq Q2	0.457568	0.116215	-0.02861	0.185512	0.115768	-0.22595	-0.10896	0.062495	-0.07984	0.274551
Peak Freq Q3	0.534699	-0.4917	-0.27032	-0.13358	-0.0055	-0.05884	0.162177	0.049635	0.015368	0.020335
Peak Freq Q4	0.476034	-0.55698	-0.21878	-0.20423	0.035996	-0.02661	0.123501	-0.01019	-0.04523	0.000665
Peak Freq SD	-0.34094	-0.53277	0.026477	-0.23672	0.003988	-0.00409	-0.11725	-0.08309	-0.08017	-0.20211
Peak to Peak	-0.32832	-0.49323	0.024035	-0.09548	-0.0818	0.014039	0.020186	-0.07239	0.126322	0.181266
Total BW Max	-0.11501	-0.27078	0.090603	0.33263	-0.199	0.083153	0.174721	-0.19268	-0.14943	0.025707
Total BW Mean	-0.02496	0.278074	-0.22718	-0.28947	-0.05144	0.129516	-0.06076	-0.02536	-0.27157	0.065472

Periodic	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Total BW Medium	-0.29348	-0.43191	-0.01692	0.518674	-0.05694	-0.06072	0.277183	0.020453	-0.09753	0.012147
Total BW Q0	0.608915	0.05727	0.0869	0.204509	0.118384	-0.02403	-0.08248	-0.11179	-0.10289	-0.16656
Total BW Q1	0.228063	-0.20836	0.046801	0.443849	-0.53585	0.112167	-0.08244	-0.49507	-0.06241	0.007049
Total BW Q2	0.31378	0.194255	0.005751	-0.20817	0.088864	0.00493	-0.36932	-0.14455	0.020107	-0.01184
Total BW Q3	0.246861	-0.15968	0.033612	0.388145	-0.53119	0.097026	-0.11436	-0.47317	-0.02986	0.00806
Total BW Q4	0.150426	-0.26627	0.072491	0.487855	-0.54221	0.121937	-0.01025	-0.49054	-0.09813	0.019329
Total BW SD	-0.23507	-0.4191	0.021483	0.515132	-0.12429	-0.06253	0.291907	0.010826	-0.10702	-0.00324

Appendix VI. Phylogenetic PCA loadings for 'mixed' calls

Mixed	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
# Peaks Max	-0.46989	-0.02541	0.265249	0.223541	0.024985	0.278691	0.106042	0.040011	0.016782	0.201818
# Peaks Mean	-0.84293	-0.21903	0.053414	0.036775	-0.22155	-0.13159	-0.01211	0.103349	0.041448	-0.04181
# Peaks Med	0.018341	-0.08785	0.193937	0.389009	-0.46872	-0.02026	0.004221	0.17954	0.022841	0.264861
# Peaks Q0	-0.20679	0.055137	-0.07211	0.460283	0.148857	-0.08118	-0.23682	-0.42368	0.321201	0.184876
# Peaks Q1	-0.56551	-0.16833	0.278427	0.283791	0.076001	-0.17235	0.175505	0.063701	-0.06957	-0.09004
# Peaks Q2	-0.61566	-0.21798	0.232588	0.005933	-0.1182	-0.03275	0.007126	0.090608	-0.05733	-0.04587
# Peaks Q3	-0.57527	-0.14577	0.015603	-0.26766	-0.38099	-0.09032	-0.09858	0.023028	0.032788	-0.10689
# Peaks Q4	-0.42585	-0.09014	-0.41822	0.010853	-0.22753	0.064562	0.057231	0.118381	0.108524	0.043675
% Nonlinear	-0.28607	-0.05391	0.034216	-0.12492	0.159311	0.134052	0.033629	-0.19127	-0.27065	0.002291
1st Freq Peak Max	0.424711	-0.43455	-0.19532	-0.09318	0.05978	-0.17305	-0.01262	0.106306	0.016817	-0.25685
1st Freq Peak Mean	0.54705	-0.48286	-0.31722	0.042834	0.204222	-0.14875	0.11161	0.202159	-0.10096	-0.00891
1st Freq Peak Med	0.122421	-0.51963	-0.17707	-0.14635	0.252393	0.042672	0.072118	-0.10888	0.009448	-0.21098
1st Freq Peak Q0	0.240955	-0.18644	-0.10484	-0.04815	0.093181	-0.03941	0.008872	0.38402	-0.29703	-0.07061
1st Freq Peak Q2	0.369867	-0.37007	-0.28977	0.065253	0.148258	-0.20093	0.086973	0.117849	-0.06725	0.070203
1st Freq Peak Q3	0.409879	-0.3942	-0.26631	0.095007	0.236404	-0.11093	0.121817	0.172588	-0.06169	0.028244
1st Freq Peak Q4	0.223149	-0.28671	-0.10168	0.043274	0.083086	-0.00723	0.144408	0.12954	-0.14961	0.075195
1st Freq Peak SD	0.019177	-0.20204	0.060679	-0.25905	0.003855	0.00578	-0.18835	-0.15526	0.278751	-0.26156
BW Max	-0.33002	-0.10749	0.034903	-0.18529	0.186746	0.239774	0.189504	-0.10127	-0.09169	0.062603
BW Mean	-0.26345	-0.04314	0.200258	-0.17099	0.315053	0.443526	0.146894	0.003067	-0.02379	0.146422
BW Med	-0.67538	-0.23734	0.129806	-0.17419	0.251935	0.169915	0.059714	-0.25183	-0.09901	-0.1311
BW Q0	-0.15856	-0.01539	-0.04891	0.286143	0.234118	0.164978	-0.10723	-0.12528	0.348115	0.03106
BW Q1	-0.31858	-0.14789	0.164621	0.232397	0.295818	-0.04326	0.24695	-0.057	-0.08355	-0.18909
BW Q2	-0.25941	-0.1669	0.153271	-0.30676	0.204076	0.134299	0.007511	-0.02481	-0.1221	0.020518
BW Q3	-0.27903	-0.08498	-0.11702	-0.42086	-0.21789	0.068298	-0.30181	-0.14512	-0.15346	-0.1668
BW Q4	-0.22553	-0.0179	-0.26721	-0.01469	-0.09605	0.251946	0.128773	0.168045	-0.01222	0.008315
BW SD	-0.51078	-0.08989	0.006328	-0.30148	0.306069	0.185168	0.096238	-0.1835	-0.12303	-0.03948
Dist to Max	0.045496	-0.01679	-0.20681	0.255496	0.044157	-0.19085	0.157601	-0.54335	-0.27034	-0.10958
Duration	0.204066	-0.0832	-0.22447	0.052593	0.064044	-0.23595	0.05596	-0.58225	-0.21584	-0.25165

Mixed	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Energy	0.431203	-0.06227	-0.27366	0.122273	0.076077	-0.31268	0.09161	-0.27402	-0.24955	-0.10801
Entropy Max	-0.75225	0.002332	0.106831	0.111575	0.195519	0.083013	0.223078	0.077236	-0.03078	0.123327
Entropy Mean	-0.91814	-0.15671	-0.02051	-0.08052	-0.00309	-0.08686	-0.03996	0.065527	0.063166	-0.05614
Entropy Med	-0.89266	-0.19751	-0.02552	-0.04644	0.038077	-0.07245	-0.00355	0.021472	-0.0416	-0.07997
Entropy Q0	-0.30519	0.000454	-0.10684	0.447281	0.144845	-0.10869	-0.09677	-0.41236	0.431622	0.175114
Entropy Q1	-0.6443	-0.139	0.193142	0.26777	0.245576	-0.1417	0.184205	0.035566	-0.05491	-0.09632
Entropy Q2	-0.6845	-0.21602	0.108828	-0.16039	0.180665	-0.03372	-0.05782	0.051259	-0.05214	-0.09048
Entropy Q3	-0.66252	-0.10108	-0.05644	-0.429	-0.20654	-0.06856	-0.20284	0.051182	0.005017	-0.14156
Entropy Q4	-0.46501	-0.08729	-0.40682	-0.01818	-0.18683	0.096682	0.017639	0.169624	0.111049	0.042003
Entropy SD	0.385912	0.176898	-0.40157	0.051614	-0.29042	-0.29947	-0.14253	-0.12163	0.124552	-0.21907
HNR Max	0.283375	0.01274	0.205344	0.065369	-0.1648	-0.09457	-0.06419	-0.00084	-0.04664	-0.24882
HNR Mean	0.48643	0.227386	0.341949	0.265163	-0.33861	0.251733	0.14165	-0.06907	-0.10303	-0.29219
HNR Med	0.232096	0.201626	0.218485	0.261263	-0.2166	0.193881	0.116923	-0.00878	0.01747	-0.2982
HNR Q0	0.222123	0.063433	0.037819	-0.07707	-0.17882	0.308484	0.10355	-0.0368	-0.25546	-0.13456
HNR Q1	0.233372	0.168904	0.102057	-0.00865	-0.26639	0.178493	0.05541	-0.07392	0.078488	-0.20159
HNR Q2	0.29471	0.166403	0.217974	0.289991	-0.23351	0.216554	0.111725	-0.08148	-0.06747	-0.17475
HNR Q3	0.234014	0.08386	0.272051	0.312798	-0.11553	0.019561	0.255993	-0.00989	-0.20618	-0.04576
HNR Q4	0.184125	0.085802	0.288998	0.091604	-0.15021	0.076446	0.045886	-0.08645	-0.17492	-0.24098
HNR SD	-0.19321	-0.18583	-0.34557	-0.32737	0.261665	-0.35016	-0.1191	-0.06623	0.033705	0.147023
Max Freq Max	0.139453	-0.76403	0.116967	-0.08306	-0.10837	0.091909	0.034878	-0.14474	0.181267	-0.08388
Max Freq Mean	0.083328	-0.93296	0.038868	0.056448	-0.06511	0.091327	0.01341	-0.03593	-0.0145	0.056867
Max Freq Med	-0.54317	-0.49532	0.123806	-0.1095	0.250002	0.116903	0.082022	-0.19294	-0.0959	-0.08451
Max Freq Q0	-0.02961	-0.30406	-0.14551	0.477092	0.259495	0.331168	-0.59593	0.112376	-0.1464	-0.05509
Max Freq Q1	-0.07561	-0.64812	0.263896	0.176609	0.208579	-0.03568	0.076217	0.101342	0.221003	-0.3469
Max Freq Q2	0.004733	-0.73486	0.209287	-0.18128	0.014647	-0.04684	0.009011	-0.00661	-0.10935	0.188248
Max Freq Q3	-0.11823	-0.61114	-0.10882	-0.17325	-0.38541	0.106897	-0.07337	-0.24838	-0.19887	0.210675
Max Freq Q4	-0.1297	-0.14874	-0.64965	0.095357	-0.0386	0.459433	0.381558	0.101965	0.114712	-0.08979
Max Freq SD	-0.65727	0.085861	-0.22493	0.136076	-0.26993	0.09706	-0.18551	-0.18586	-0.06948	-0.11854
Min Freq Max	0.230331	-0.74464	0.108922	-0.03404	-0.16034	0.028357	-0.0157	-0.11927	0.208387	-0.10213

Mixed	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Min Freq Mean	0.118345	-0.92635	-0.00512	0.083568	-0.1346	-0.0076	-0.02074	-0.04174	-0.01754	0.013723
Min Freq Med	0.514921	-0.48662	-0.05659	0.20012	-0.08803	-0.17319	0.029052	0.212911	0.039713	0.145597
Min Freq Q0	0.058749	-0.328	-0.13298	0.363341	0.152098	0.271741	-0.59834	0.197013	-0.36346	-0.07907
Min Freq Q1	0.153392	-0.63403	0.179112	0.027534	0.016727	-0.0084	-0.09873	0.15966	0.317387	-0.25633
Min Freq Q2	0.17665	-0.71909	0.134877	-0.00186	-0.11827	-0.14158	0.005158	0.008942	-0.0428	0.198961
Min Freq Q3	0.072913	-0.64473	-0.0382	0.115748	-0.28302	0.072647	0.142061	-0.17883	-0.11535	0.36973
Min Freq Q4	-0.00828	-0.1655	-0.5997	0.12298	0.016277	0.383449	0.370282	0.012454	0.144401	-0.11221
Min Freq SD	-0.68049	0.117139	-0.22052	0.148043	-0.22247	0.112923	-0.17606	-0.16963	-0.06763	-0.0789
Peak Freq Max	0.194556	-0.7536	0.119201	-0.04998	-0.13364	0.056147	0.011985	-0.13216	0.19607	-0.09401
Peak Freq Mean	0.109256	-0.93299	0.011656	0.066707	-0.10062	0.043911	-0.00709	-0.04219	-0.0112	0.037864
Peak Freq Med	0.203439	-0.79407	0.094095	-0.08482	-0.01742	0.03876	0.041929	-0.04869	0.081037	-0.03245
Peak Freq Q0	0.024444	-0.33849	-0.11925	0.417929	0.214698	0.273365	-0.62092	0.18018	-0.28335	-0.06561
Peak Freq Q1	0.053908	-0.67264	0.241915	0.076692	0.123624	-0.02114	-0.03069	0.12368	0.254382	-0.34429
Peak Freq Q2	0.116579	-0.7546	0.155995	-0.09367	-0.05465	-0.0993	0.022555	0.012566	-0.09491	0.200252
Peak Freq Q3	-0.03508	-0.65128	-0.10141	-0.02977	-0.39896	0.094813	-0.01011	-0.24278	-0.15221	0.26686
Peak Freq Q4	-0.01711	-0.15089	-0.6474	0.114145	-0.0208	0.43742	0.375456	0.052945	0.126699	-0.12129
Peak Freq SD	-0.68186	0.106625	-0.20403	0.153304	-0.24209	0.082021	-0.17051	-0.16472	-0.08528	-0.10981
Peak to Peak	0.263082	-0.09672	-0.3252	0.091216	0.159762	-0.35524	0.123092	-0.1307	-0.29089	-0.10664
Total BW Max	-0.4181	-0.05191	0.064196	0.316147	-0.09155	-0.2553	0.225583	0.035633	-0.10071	0.016748
Total BW Mean	-0.58129	-0.03185	0.115322	0.380533	-0.19184	-0.17341	0.254505	0.049436	-0.13188	0.05433
Total BW Medium	-0.69174	-0.1627	-0.05508	0.299572	-0.10946	-0.18689	0.141549	0.150347	-0.13931	-0.08453
Total BW Q0	-0.29815	-0.03213	-0.00879	0.354919	0.097134	-0.14309	-0.09624	-0.16628	0.218831	0.172565
Total BW Q1	-0.44054	-0.08653	0.001833	0.425941	0.061163	-0.16921	-0.00271	0.121387	-0.08195	-0.01696
Total BW Q2	-0.50138	-0.08295	-0.01373	0.056243	0.037904	-0.36	0.204371	0.141768	0.012331	-0.11393
Total BW Q3	-0.52848	-0.08191	-0.0736	-0.17484	-0.29839	-0.24139	0.006839	0.21751	-0.09689	-0.13119
Total BW Q4	-0.52256	-0.06479	-0.27409	0.039998	-0.13678	-0.03373	-0.02598	0.237288	0.105381	-0.00202
Total BW SD	-0.28546	0.000705	-0.29084	-0.01101	-0.0741	-0.36052	-0.09361	0.152381	0.044089	-0.20186

Appendix VII. List of linear mixed effects models retained during model selection (using AICc criteria) by call type. Individual animal was the random effect, and when retained during model selection date was nested within individual. Models are listed using R notation.

Noisy	
PC1	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities
PC2	(1|Animal/Date)+Age*Sex
PC3	(1|Animal/Date)+Age*Sex+SNR+Nonlinearities
PC4	(1|Animal/Date)+Age*Sex+SNR)
PC5	(1|Animal/Date)+Age*Sex+SNR)
PC6	(1|Animal/Date)+Age*Sex+SNR+Nonlinearities
PC7	(1|Animal/Date)+Age*Sex*SNR
PC8	(1|Animal/Date)+Age*Sex
PC9	(1|Animal/Date)+Age*Sex*SNR)
PC10	(1|Animal)+Age*Sex

Periodic	
PC1	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities
PC2	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities
PC3	(1|Animal/Date)+Age*Sex*SNR
PC4	(1|Animal/Date)+Age*Sex
PC5	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities
PC6	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities
PC7	(1|Animal/Date)+Age*Sex+SNR+Nonlinearities
PC8	(1|Animal/Date)+Age*Sex+SNR
PC9	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities
PC10	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities

Mixed	
PC1	(1|Animal/Date)+Age*Sex+SNR+Nonlinearities
PC2	(1|Animal)+Age*Sex*SNR+Nonlinearities
PC3	(1|Animal)+Age*Sex+SNR+Nonlinearities
PC4	(1|Animal)+Age+Sex
PC5	(1|Animal)+Age+Sex
PC6	(1|Animal)+Age*Sex+SNR
PC7	(1|Animal)+Age+Sex
PC8	(1|Animal/Date)+Age*Sex*SNR+Nonlinearities
PC9	(1|Animal)+Age*Sex+SNR+Nonlinearities
PC10	(1|Animal)+Age*Sex*SNR

Appendix VIII. Linear mixed effects models for 'noisy' calls showing how PCA components varied with age, sex, signal to noise ratio (SNR) and presence/absence of nonlinearities with random effects of date was nested within individual. Covariates of age, sex and their interaction are shown in table 8. For sex, males were referential in the model. For nonlinearities, absence was referential. Significance (p < 0.05) is shown in bold.

PC	SNR				Age:SNR				Sex:SNR			
	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	-0.296	-2.047	1.444	0.7427	0.824	0.120	1.533	**0.0241**	0.123	-2.109	2.457	0.9158
PC2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PC3	-0.769	-0.994	-0.512	**0.0000**	NA	NA	NA	NA	NA	NA	NA	NA
PC4	-1.082	-1.268	-0.888	**<2e-16**	NA	NA	NA	NA	NA	NA	NA	NA
PC5	-0.749	-0.95	-0.552	**0.0000**	NA	NA	NA	NA	NA	NA	NA	NA
PC6	-0.512	-0.695	-0.31	**0.0000**	NA	NA	NA	NA	NA	NA	NA	NA
PC7	0.275	-0.324	0.855	0.3631	-0.466	-0.702	-0.214	**0.0002**	0.086	-0.688	0.861	0.8290
PC8	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PC9	0.704	0.178	1.241	**0.0098**	-0.361	-0.58	-0.143	**0.0013**	-0.803	-1.491	-0.085	**0.0259**
PC10	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

PC	Age:Sex:SNR				Nonlinearities			
	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	0.078	-1.066	1.165	0.8907	-3.809	-5.786	-1.829	**0.0002**
PC2	NA	NA	NA	NA	NA	NA	NA	NA
PC3	NA	NA	NA	NA	0.989	0.092	1.8950	**0.0318**
PC4	NA	NA	NA	NA	NA	NA	NA	NA
PC5	NA	NA	NA	NA	NA	NA	NA	NA
PC6	NA	NA	NA	NA	1.257	0.559	1.987	**0.0006**
PC7	-0.073	-0.454	0.308	0.7130	NA	NA	NA	NA
PC8	NA	NA	NA	NA	NA	NA	NA	NA
PC9	0.084	-0.277	0.426	0.6391	NA	NA	NA	NA
PC10	NA	NA	NA	NA	NA	NA	NA	NA

Appendix IX. Linear mixed effects models for 'periodic' calls showing how PCA components varied with age, sex, signal to noise ratio (SNR) and presence/absence of nonlinearities with random effects of date was nested within individual. Covariates of age, sex and their interaction are shown in table 8. For sex, males were referential in the model. For nonlinearities, absence was referential. Significance (p < 0.05) is shown in bold.

PC	SNR				Age:SNR				Sex:SNR			
	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	3.419	2.933	3.902	**<2e-16**	-0.494	-0.693	-0.302	**7.76E-07**	-0.012	-0.842	0.816	0.977
PC2	0.265	-0.281	0.814	0.3436	0.241	0.02	0.459	**0.0322**	-0.004	-0.943	0.934	0.994
PC3	0.6334	0.365	0.908	**4.85E-06**	-0.11	0.365	0.908	**0.044**	-1.749	-2.111	-1.191	**3.29E-12**
PC4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PC5	-0.465	-0.765	-0.164	**0.0025**	0.076	-0.043	0.195	0.2115	-0.532	-1.037	-0.022	**0.0412**
PC6	-0.328	-0.588	-0.082	**0.0108**	-0.216	-0.313	-0.112	**2.52E-05**	-0.007	-0.425	0.432	0.9747
PC7	-0.234	-0.335	-0.121	**9.33E-06**	NA	NA	NA	NA	NA	NA	NA	NA
PC8	0.156	0.058	0.251	**0.0016**	NA	NA	NA	NA	NA	NA	NA	NA
PC9	0.689	0.439	0.929	**2.87E-08**	-0.339	-0.433	-0.236	**7.59E-12**	-0.424	-0.839	-0.019	**0.0429**
PC10	-0.356	-0.588	-0.123	**0.0027**	-0.165	-0.256	-0.073	**0.0004**	0.324	-0.055	0.73	0.1055

PC	Age:Sex:SNR				Nonlinearities			
	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	-0.321	-0.854	0.203	0.235	1.125	0.573	1.681	**7.32E-05**
PC2	-0.51	-1.11	0.087	0.0936	-0.541	-1.163	0.081	0.0894
PC3	0.943	0.654	1.231	**2.31E-10**	NA	NA	NA	NA
PC4	NA	NA	NA	NA	NA	NA	NA	NA
PC5	0.526	0.214	0.839	**0.0011**	1.008	0.661	1.346	**9.77E-09**
PC6	-0.04	-0.305	0.224	0.7654	-2.313	-0.305	0.224	0.7654
PC7	NA	NA	NA	NA	-1.617	-1.93	-1.327	**<2e-16**
PC8	NA	NA	NA	NA	NA	NA	NA	NA
PC9	0.484	0.225	0.731	**0.0001**	-1.468	-1.753	-1.201	**0.0002**
PC10	-0.193	-0.429	0.051	0.1171	2.026	-0.429	0.051	**<2e-16**

Appendix X. Linear mixed effects models for 'mixed' calls showing how PCA components varied with age, sex, signal to noise ratio (SNR) and presence/absence of nonlinearities with random effects of date was nested within individual. Covariates of age, sex and their interaction are shown in table 8. For sex, males were referential in the model. For nonlinearities, absence was referential. Significance (p < 0.05) is shown in bold.

PC	SNR				Age:SNR				Sex:SNR			
	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	2.914	2.294	3.558	**<2e-16**	NA	NA	NA	NA	NA	NA	NA	NA
PC2	-1.781	-2.938	-0.672	**0.0024**	0.659	0.235	1.144	**0.0048**	3.427	0.822	5.689	**0.0076**
PC3	-0.307	-0.589	-0.001	**0.0416**	NA	NA	NA	NA	NA	NA	NA	NA
PC4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PC5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PC6	-0.438	-0.686	-0.183	**0.0008**	NA	NA	NA	NA	NA	NA	NA	NA
PC7	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
PC8	0.839	0.028	1.523	**0.0288**	-0.471	-0.724	-0.149	**0.0008**	-0.504	-1.647	0.782	0.4371
PC9	-0.202	-0.387	-0.015	**0.0376**	NA	NA	NA	NA	NA	NA	NA	NA
PC10	0.972	0.417	1.45	**0.0003**	-0.427	-0.592	-0.207	**3.40E-05**	-1.105	-2.158	-0.388	**0.0288**

PC	Age:Sex:SNR				Nonlinearities			
	Coeff	2.5%	97.5%	P	Coeff	2.5%	97.5%	P
PC1	NA	NA	NA	NA	-5.45	-10.398	-0.623	**0.0302**
PC2	-1.76	-3.218	-0.109	**0.0344**	-0.58	-5.411	4.007	0.8096
PC3	NA	NA	NA	NA	-0.034	-3.046	2.939	0.9824
PC4	NA	NA	NA	NA	NA	NA	NA	NA
PC5	NA	NA	NA	NA	NA	NA	NA	NA
PC6	NA	NA	NA	NA	NA	NA	NA	NA
PC7	NA	NA	NA	NA	NA	NA	NA	NA
PC8	0.003	-0.759	0.703	0.9937	-0.186	-2.74	2.084	0.8798
PC9	NA	NA	NA	NA	-2.07	-4.225	0.091	0.064
PC10	0.474	0.008	1.116	0.1306	NA	NA	NA	NA

Appendix XI. Linear mixed effects models for 'noisy' PCA component 2 varied with age (in days), sex, signal to noise ratio (SNR) and categorical classification of nonlinearities. Definitions for nonlinearity classes, including chaotic ('Ch'), subharmonics ('Su') and limit cycles (aka harmonics, 'Ha') are shown in table 2. Date was nested within individual animal as a random effect. Significant (p < 0.05) variables are shown in bold.

PC	Intercept			
		CI		
	Coeff	2.5%	97.5%	P
Intercept	-0.472	0.079	4.889	0.6623
Age	-0.006	0.476	2.076	0.9882
Sex	0.735	0.029	11.576	0.7350
Age:Sex	0.106	0.307	4.029	0.8730
Nonlin. Ch Su	0.679	0.019	197.580	0.7727
Nonlin. Ha	2.055	3.330	18.865	**0.0000**
Nonlin. Ha Ch	0.014	0.225	4.582	0.9851
Nonlin. Ha Ch Su	1.596	0.153	158.689	0.3679
Nonlin. Ha Su	0.825	0.265	19.652	0.4526
Nonlin. Ch	1.432	1.176	14.903	**0.027**

Appendix XII. Example recordings of Zola's call types, growls and moans. Audio files are attached in the supplemental material CD.

Appendix XIII. Example recordings of Zola's trials, showing correct and incorrect responses to growl and moan stimuli. Audio files are attached in the supplemental material CD.

Appendix XIV. Models used to calculate model average.

Model	DF	Log Likelihood	AICc	Delta	Weight
1	5	-231.34	472.76	0.00	0.58
2	5	-231.81	473.70	0.94	0.36
3	10	-229.65	479.62	6.85	0.02
4	4	-236.19	480.45	7.68	0.01
5	5	-235.31	480.71	7.95	0.01
6	5	-235.62	481.32	8.56	0.01
7	6	-234.74	481.60	8.83	0.01
8	5	-236.19	482.47	9.70	0.00

Appendix XV. Table of factor loadings from the factor analysis for mixed data (FAMD), separated by call type (growls versus moans), and by individual (individual animal differences for (growls and moans). When separated by call type, sounds were clustered across three components. However, when separated by individual calls were clustered best with the first two components. Thus, only two dimensions are shows for calls clustered by individual. Parameters correspond with the definitions shown in Table 1*.

Growls vs. Moans

Continuous variables

	Dim. 1	Dim. 2	Dim. 3
Dur	0.363	-0.208	0.445
Peak.F	0.535	0.429	-0.197
Beg. F	0.451	0.464	0.369
End.F	0.402	0.275	0.038
Mid.F	0.544	0.433	0.033
Max.F	-0.561	0.496	0.224
Wiener	-0.686	0.387	0.189
HNR	0.698	-0.286	0.013

Categorical variables

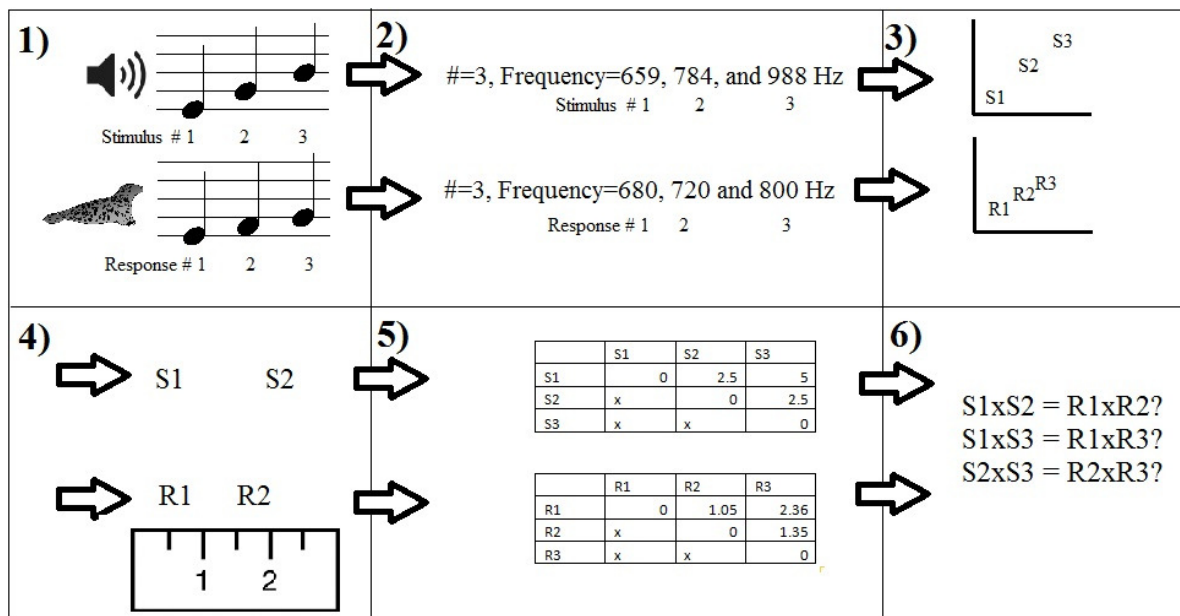|	Dim. 1	Dim. 2	Dim. 3
FF_A	-2.083	-0.131	-0.213
FF_P	2.083	0.131	0.213
SR_0	-2.083	-0.131	-0.213
SR_1-3	2.083	0.131	0.213

Moans by Individual

Continuous variables

	Dim. 1	Dim. 2
Dur	0.11	-0.263
Peak.F	0.034	0.265
Beg. F	0.631	0.664
End.F	0.515	0.532
Mid.F	0.538	0.494
Max.F	-0.338	0.559
Wiener	-0.1	0.553
HNR	0.204	0.141
FF	-0.793	0.273
SR	0.134	0.553

Growls by Individual

Continuous variables

	Dim. 1	Dim. 2
Dur	0.01	-0.393
Peak.F	0.149	0.292
Beg. F	0.277	0.634
End.F	0.768	0.201
Mid.F	0.707	0.294
Max.F	-0.042	0.624
Wiener	0.44	-0.175
HNR	-0.286	0.205

*Parameter abbreviations: Dur=duration, Peak.F= peak frequency at maximum amplitude of call, Beg.F= peak frequency at beginning of call, End.F= peak frequency at end of call, Mid.F = peak frequency at middle of call, Max.F= maximum frequency reached at any point within call, Wiener= wiener entropy, HNR= harmonic to noise ratio, FF= fundamental frequency (if categorical absent (A) versus present (P)), SR= signal to noise ratio.

Appendix XVI. Example recordings of Zola's song imitation trials. Two examples are provided, corresponding to Figures 5.5 and 5.6. The first 'song' was the first 5 'notes' of 'Mary had a little lamb', showing first the played signal, and then the seals call in response. The second 'song' was the complete 7 'note' "twinkle twinkle little star", showing first the played signal, and then the seals call in response. An additional video example is provided of the complete 7 'note' song "starwars", showing first the played signal, and then the seals response. Audio and video files are attached in the supplemental material CD.

Appendix XVII. Illustration of how the statistical similarity between calls was measured (diagram shown on the next page). 1) The sound stimuli played and the seal's responses were recorded. In this case a 3 'note' stimulus was played, and the seal responded with a 3 'note' call. Throughout the analysis, responses were directly compared to their stimulus counterpart, such that stimulus 1 was directly compared to response 1, stimulus 2 to response 2, etc.. 2) Acoustic parameters were measured for each 'note' individually. 3) Using the measured acoustic parameters, signals and responses were plotted (separately) using a cluster analysis technique designed to measure dissimilarity ('Daisy'). Rather than using fixed measurement units (such as Hz for frequency), sounds were plotted onto a new scale based on the variability observed across all sounds (in this example only three 'notes' are shown, but during the analysis all of the sounds were plotted, thus providing a measure of overall variability). 4) The distance ('Gower') between points was used as a measure of similarity between calls. The distance between all signals was measured separately from the distance between all calls. 5) The measured distances were combined into distance matrices. Separate matrices were produced for the signals and the calls. For example, the vector S1 x S1 represents the distance between stimulus one to stimulus one. It is the same sound, and thus there is no difference in distance. For the vector S1 x S2, the distance between stimulus one to stimulus two was measured. Smaller values indicate the signals are more similar, while larger values indicate they are less similar. 6) The matrices were lined up such that S1 corresponds to the first sound played, and R1 corresponds to the first sound the seal responded with. The mantel test is then used to measure how similar the two matrices are in comparison to chance. To test this, the matrices were pseudo-randomly mixed with 1,000 permutations, such that songs (this example consisted of one 3 'note' song) were kept intact.

1) Stimulus # 1   2   3

Response # 1   2   3

2) #=3, Frequency=659, 784, and 988 Hz
Stimulus # 1   2   3

#=3, Frequency=680, 720 and 800 Hz
Response # 1   2   3

3)
S3
S2
S1

R2 R3
R1

4) S1   S2

R1   R2

1   2

5)

	S1	S2	S3
S1	0	2.5	5
S2	x	0	2.5
S3	x	x	0

	R1	R2	R3
R1	0	1.05	2.36
R2	x	0	1.35
R3	x	x	0

6)
S1xS2 = R1xR2?
S1xS3 = R1xR3?
S2xS3 = R2xR3?

Appendix XVIII. Example recording of one of Janice's test trials, corresponding to Figure 5.7. An additional sound file example of Gandalf's test trials is also provided. While our seals successfully matched formant frequencies (predominantly by varying the 2nd formant), they matched poorly in comparison to previous recordings of Hoover. An example of Hoover speaking (obtained from https://web.archive.org/web/20110506041412/http://www.st-andrews.ac.uk/~wtsf/Hoover.html) is also shown. Audio files are attached in the supplemental material CD.

Appendix XIX. Example recordings of the novel call type observed on the Isle of May. Experts who have extensive experience working with grey seals and examining their vocalisations had not previously seen this call type (personal communication Simon Moss, Susanne McCulloch, Thomas Götz). It was initially observed in an adult female seal, and upon weaning pups raised nearby began to produce this call. Audio files of the adult female and a weaned pup are attached in the supplemental material CD.