

UNIVERSITY OF CALGARY

The Acoustic Structure and Ontogeny of Vervet Monkey Vocalizations

by

Colin James Dubreuil

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN ANTHROPOLOGY

CALGARY, ALBERTA

NOVEMBER, 2019

© Colin James Dubreuil 2019

Abstract

Vervet monkey (*Chlorocebus pygerythrus*) vocalizations have been the subject of considerable research, most notably for their putative language-like qualities. While this focus has inspired a productive research effort investigating vocal communication in non-human primates, it has diverted attention away from other, non-linguistically inspired mechanisms by which vocal signals exert their effects on receivers. My research focuses on two vocal classes, grunts and alarm calls, and how their acoustic structures vary according to sender-specific attributes, including age, sex, body size, and identity - all of which have the potential to influence receiver response. I recorded calls from three wild groups of vervet monkeys over a 7-month period on the Samara Game Reserve, South Africa. I used random forest models to determine whether grunts varied in structure in relation to caller age, sex, identity and context, and whether alarms varied with sex and caller identity. I performed a cluster analysis to determine whether alarms segregated into different call types based on variation in acoustic structure. Finally, I used a series of mixed effects models to determine whether call structure in males and females correlated with overall body size (using body weight as a proxy). I found grunts varied in structure with age and sex, but not among individuals or contexts. Alarm calls varied with both sex and identity. My cluster analysis identified two qualitatively distinct alarm call types corresponding to the calls of males and females respectively. My analysis demonstrated that the relationship between body weight and the distribution of acoustic energy throughout alarm calls differed between the sexes. Relative to females, acoustic energy in the calls of males was concentrated at disproportionately lower frequencies. These results suggest that alarm calls may be under selection to exaggerate caller body size and/or extend the effective range of these signals in males. My results highlight that vocalizations can vary in structure along a number of dimensions simultaneously. While it is possible that in some instances, vocalizations convey specific information surrounding their context of production, it is likely that the mechanisms by which vocal signals exert their effects on receivers are more numerous and diverse.

Acknowledgements

Firstly, I would like to thank my supervisory team: Dr. Hugh Notman, Dr. Mary Pavelka, and Dr. Peter Henzi and Dr. Louise Barrett for their support, direction, and mentorship. I'd like to thank Dr. Kurt Hammerschmidt, and Dr. Julia Fischer for their hospitality and help with my acoustic analysis. Special thanks to Dr. Tyler Bonnell for his help and guidance with my statistical analysis, and his endless patience. Thanks to Dr. Jonathan Jarrett and Dr. Richard McFarland for their help with the weight data.

I would also like to thank the field team: Dr. Christopher Young, Miri Minkner, Rose Blersch, Christina Nord, Stephen Morris, Lucy Twitcher, Tatiana Bortolato, and Pia Marlina for their help, patience, and perseverance in the field. None of this would have been possible without you. Special thanks to Kitty and Richard Viljoen for making the field feel like home.

Financial support for this research was provided by the Natural Sciences and Engineering Research Council of Canada, the Killam Trust, the Faculty of Graduate Studies, and the Department of Anthropology at the University of Calgary.

Special thanks to Mom and Dad, for constantly being there for me. Thanks to Alex, Veronica, Lily and Theodore for all the smiles over the last 5 years. Finally, I'd like to thank Lucy Millington for being there through this entire process, whether in person or through skype. Your support and patience have meant a lot to me.

Contents

Chapter 1 Introduction	1
1.1 Theoretical Framework.....	1
1.2 Research Goals, and Layout of this Thesis	4
1.3 Literature Review	7
Source Filter Theory and the Relationship Between Body Size and Acoustic Structure	7
Variation in Relation to Body Size.....	9
Variation Related to Age and Sex.....	10
Variation Related to Individual Identity.....	14
Variation related to Motivation.....	18
Variation related to Context	20
Information and Influence	24
1.4 Brief outline of subsequent chapters	27
Chapter 2 Sources of acoustic variation in the grunt vocalizations of wild vervet monkeys	28
2.1 Introduction	28
Vervet Grunts.....	30
2.2 Methods.....	34
Study Species	34
Study Site and Population.....	36
Vocal data collection in the field.....	37
2.3 Data Analysis.....	39
Assigning Contexts to calls.....	39
Selecting calls for analysis.....	42
Acoustic analysis	44
Statistical Analysis.....	49
2.4 Results.....	51
Classification of caller sex	51
Classification by age class	54
Classification of individuals using their grunt vocalizations	58
Effects of context on grunt variation	59
2.5 Discussion.....	70
Age and Sex.....	70

Individual Identity	72
Context.....	73
Chapter 3 Sex Differences in the terrestrial alarm calls of vervet monkeys.....	82
3.1 Introduction	82
Vervet alarm calls.....	84
3.2 Methods.....	87
Vocal data collection in the field.....	87
3.3 Data Analysis.....	88
Selecting calls for analysis.....	88
Acoustic Analysis.....	90
Statistical analysis	92
3.4 Results.....	96
Alarm calls – General description	96
Acoustic differences between male and female alarm calls	97
Cluster analysis:	101
Classification by caller ID: Individual differences in alarm calls.....	106
3.5 Discussion.....	108
Intergradation between call types.....	110
Individual distinctiveness.....	112
Chapter 4 Vocal allometry in vervet monkeys.....	117
4.1 Introduction	117
4.2 Methods.....	124
Vocal data collection in the field.....	124
Acoustic analysis	125
Variable selection.....	126
Body mass	127
4.3 Data Analysis.....	129
Models	129
4.4 Results.....	131
Spectral parameters.....	131
Temporal parameters	140
4.5 Discussion.....	149
Communication with the predator	152

Sexual selection.....	155
Chapter 5 General Discussion	159
5.1 Overview	159
5.2 Summary of findings	159
Chapter 2: Acoustic variability in the structure of vervet grunts.....	159
Chapter 3: Acoustic variation in the structure of terrestrial alarm calls	160
Chapter 4: The influence of body size on the acoustic structure of grunts and terrestrial alarm calls	161
5.3 Implications for the field of primate communication.....	162
Functional reference and “arbitrary” signal structure.....	162
The informational content of signals	165
5.4 Limitations of my study, and directions for Future research.....	166
References	170
Appendices.....	196
Appendix to Chapter 2	196
Appendix to Chapter 4	200

List of Tables

Table 2.1: Group composition over the course of data collection.	37
Table 2.2: Definitions of the different contexts in which grunts were produced.	41
Table 2.3: Total number of grunts analyzed for each age sex class.....	43
Table 2.4: Variables measured from grunt vocalizations	48
Table 2.5: Confusion matrix - The number of correct classification and misattributions for each caller category (adult Males vs adult Females).	52
Table 2.6: Confusion matrix - The number of correct classifications and misattributions for calls produced by different age classes.	55
Table 2.7: Confusion matrix - The number of correct classifications and misattributions for calls produced by each individual in the analysis	59
Table 2.8: Contextual breakdown of Grunts by adult males and adult females	60
Table 2.9: Confusion matrix: The number of correct classifications and misattributions for calls produced by females across three contexts.	62
Table 2.10: Confusion matrix - The number of correct classifications and misattributions for calls produced by males across four contexts.	67
Table 3.1: Acoustic parameters used in random forest models.....	91
Table 3.2: Transformations applied to the raw data of acoustic parameters before scaling data for cluster analysis.....	95
Table 3.3: Confusion matrix: The number of correct classification and misattributions for each sex (adult males vs adult female).	98
Table 3.4: Distribution of males and females between the two clusters	102
Table 3.5: Confusion matrix - The number of correct classification and misattributions for each individual	107
Table 4.1: Number of calls used in the analysis broken down by age/sex class.....	124
Table 4.2: Acoustic Variables used in the analysis.....	127

List of Figures

Figure 2.1: Spectrograms representing grunts A) produced by females, B) produced by males.	44
Figure 2.2: Variable importance plot for predictor variables from random forests classifications used for predicting caller sex.	53
Figure 2.3: Boxplots representing the distribution of four variables which varied significantly between adult males and adult females.	54
Figure 2.4: Variable importance plot for predictor variables from random forests classifications used for predicting caller age category.	56
Figure 2.5: Boxplots representing the distribution of four variables which varied significantly among age classes.	57
Figure 3.1: Oscillograms and spectrograms representing typical alarm calls produced by females and males to terrestrial predators.	89
Figure 3.2: Variable importance plot for predictor variables from random forest classifications used for predicting caller sex.	99
Figure 3.3: Boxplots representing the distribution of four variables which vary significantly between adult males and adult females.	100
Figure 3.4: Silhouette values for the cluster analysis as a function of the number of clusters.	102
Figure 3.5: Results of the soft cluster analysis: Scatterplot represents the calls of adult males and adult females plotted against the first two principal components.	103
Figure 3.6: : Histogram representing cluster membership.	104
Figure 3.7: Typical and a-typical alarm calls produced by males and females, as highlighted by the cluster analysis.	105
Figure 3.8: Variable importance plot for predictor variables from random forests classifications used for predicting caller identity.	108
Figure 4.1: Oscillograms and spectrograms representing Grunts and Alarm calls produced by males and females.	126
Figure 4.2: Weight data used to model the interaction between body weight and each acoustic parameter.	129
Figure 4.3: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and DF1 in grunts and alarms produced by males and females, and (B) the relationship between body weight and variability in DF1 in grunts and alarms produced by males and females.	132
Figure 4.4: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and DFA2 in grunts and alarms produced by males and females, and (B) the	

relationship between body weight and variability in DFA2 in grunts and alarms produced by males and females.....	132
Figure 4.5: Density plots represent the rate of change in DF1 as body size increased (Hz/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms).....	134
Figure 4.6: Scatterplots illustrating the relationship between body mass and DF1 for alarm calls and grunts.	135
Figure 4.7: Density plots represent the rate of change in the variability of DF1 as body size increased (Hz/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms).	136
Figure 4.8: Density plots represent the rate of change in DFA2 as body size increased (Hz/kg) predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms).....	138
Figure 4.9: Scatterplots illustrating the relationship between body weight and DFA2 for alarm calls and grunts.	139
Figure 4.10: Density plots represent the rate of change in the variability of DFA2 as body size increased (Hz/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms).	140
Figure 4.11: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and mean element length in grunts and alarms produced by males and females, and (B) the relationship between body weight and variability in the mean element length in grunts and alarms produced by males and females.	141
Figure 4.12: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and call duration in grunts and alarms produced by males and females, and (B) the relationship between body weight and variability in call duration in grunts and alarms produced by males and females.	141
Figure 4.13: Density plots represent the relationship between mean element length and body size (ms/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms).	143
Figure 4.14: Scatterplots illustrating the relationship between body weight and mean element length for alarm calls and grunts.....	144
Figure 4.15: Density plots representing the relationship between variability in element duration and body size, as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms).	145
Figure 4.16: Density plots representing the relationship between call duration and body size (ms/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). .	147
Figure 4.17: Scatterplots illustrating the relationship between body weight and call duration for alarm calls and grunts.	148

Figure 4.18: Density plots representing the relationship between variability in call duration and body size, as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). 149

Chapter 1 Introduction

1.1 Theoretical Framework

A biological signal is defined as any anatomical trait or behaviour that has evolved to influence the behaviour of other individuals in ways that are beneficial for the signaler, and that is effective because receiver responses have evolved as well (Smith and Harper 2003, Scott-Phillips 2008). Researchers interested in the communication systems of non-human animals are compelled by two over-arching and interrelated questions regarding signals: how is signal variation related to function, and what are the mechanisms by which signals exert their effects on receivers (Garcia and Favaro 2017). Classic ethological approaches assume that signal structure is inextricably related to signal function, and by extension, that structural variation is constrained by selection.

The relationship between the physical structure of vocal signals and their purported function has been addressed in a number of empirical studies. These investigations have explored the relationship between acoustic structure and function, as well as the mechanism(s) by which the former affects the latter. Signal structure has been found to vary in relation to a number of factors, including (but not limited to) the distance between callers and receivers (Bradbury and Vehrencamp 2011), habitat type (Boncoraglio and Saino 2007), caller motivation (Manser 2010, Briefer 2012), identity (Tibbetts and Dale 2007), sex (Snowdon 2004), age (Ey et al. 2007a), body size (Fitch 2006), and context (Townsend and Manser 2013). From an evolutionary (functional) standpoint, researchers have put forth a number of explanatory frameworks to account for this variation, including sexual selection (Puts et al. 2016, Zuberbühler 2006), predator avoidance (Caro 2005, Isbell and Bidner 2016), kin selection (Zuberbühler 2006), the maintenance of group cohesion (Zuberbühler and Fedurek 2016, Gall and Manser 2017), and the regulation of social dynamics (Zuberbühler and Fedurek 2016, Gustison et al. 2019). The mechanisms by which

signals are proposed to exert their effects on receivers are equally diverse, and include a range of possible ways that signals can influence receiver behaviour, including the direct effects of the acoustic energy on a receiver's emotional or physiological state (Owren and Rendall 2001), the transfer of individually or context specific information (Seyfarth et al. 2010), or via the effects of learning (Owren and Rendall 1997). While these continue to present productive avenues of exploration, their diversity and multiplicity also highlight the obvious fact that relationships between signal structure, function and mechanism are varied and complex.

Because of the relatively close evolutionary relationship between non-human primates (hereafter, primates) and humans, researchers interested in primate vocal communication have typically emphasized the possible continuities between primate vocal signaling behaviour and human language (Fischer 2017, Liebal et al. 2014). Interest in this field owes a great deal to now classic studies that showed that vervet monkeys (*Chlorocebus pygerythrus*) produce three predator-specific alarm call types, and that receivers respond to each call type with what appeared to be functionally adaptive escape behaviours appropriate to the class of predator that originally elicited the signal (Struhsaker 1967a; Seyfarth et al. 1980a,b, described in more detail latter in this chapter). Because these vocalizations seemed to refer to specific objects or events in a caller's environment, these findings were initially interpreted as representing a possible homologous link between the vocalizations of non-human primates and words used in human language. Additionally, these calls seemed to be arbitrary in structure relative to their purported referent, in much the same way as the acoustic structure of human words are typically arbitrary with respect to their meaning. Numerous other examples of animal signals that appear to convey specific information about their context of production were subsequently reported across a wide variety of taxa (e.g., Chickens, *Gallus gallus*: Evans et al. 1993, Evans and Marler 1994, Evans and Evans

1999, Vervet monkeys: Cheney and Seyfarth 1982A, Seyfarth and Cheney 1984; ringtailed lemurs, *Lemur catta*: Macedonia 1990; rhesus macaques, *Macaca mulatta*: Hauser and Marler 1993; Hauser 1998; Diana monkeys, *Cercopithecus diana*: Zuberbühler et al 1999; Ravens, *Corvus corax*: Bugnyar et al. 2001; Meerkat, *Suricata suricatta*: Manser 2001, Manser et al. 2001; Gunnison's prairie dogs, *Cynomys gunnisoni*: Kiriazis and Slobodchikoff 2006, Slobodchikoff and Placer 2006; yellow warblers *Dendroica petechia*: Gill and Sealy 2004; black-capped chickadees, *Poecile atricapilla*: Templeton et al. 2005; Capuchin monkeys; Di Bitetti 2003; Wheeler 2010, dogs, *Canis familiaris*: Faragó et al. 2010). These observations have since generated a great deal of discussion surrounding the mechanisms by which signals affect receiver behaviour, and how cognition may (or may not) be integral to the process.

Mechanistically, it is thought that receivers are able to infer the context in which these “functionally referential” signals are produced because they exhibit a high degree of ‘stimulus specificity’ (Macedonia and Evans 1993), meaning there is a predictable, statistical relationship between a particular signal type and a given external event. This pairing is thought to allow receivers to extract meaningful “information” about their external world due to the probabilistic association between a specific event and a particular call type (Wheeler and Fischer 2012; Seyfarth et al. 2010; Scarantino and Clay 2015). To determine whether receivers are able to extract this specific information from these signals, researchers generally measure their responses to a given call type, usually in the absence of the actual eliciting stimulus, or any potential cues relating to the call's usual context of production (usually through the use of playback experiments). The logic of this approach is that if signals effectively convey information about the external world, receivers should respond with contextually appropriate behaviours, even in the absence of these disambiguating cues (a criterion called ‘context independence’; Macedonia and Evans 1993).

Assuming that these signals exert their effects on receivers is through a consistent and close association between the production of a call and some specific context, the mechanisms by which functionally referential signals are thought to exert their effects on receivers are fundamentally different from those of many other signal types; while signal structure is generally seen as being *adapted* to influence receiver behaviour in a specific way, researchers have often highlighted the fact that the structure of functionally referential signals are actually *arbitrary* with regards to their purported referent (Morton 2017). Consequentially, the strong focus on functional reference has come under scrutiny by several researchers (Owings and Morton 1998, Owren and Rendall 2001, Wheeler and Fischer 2012, 2015). One common criticism of the “functionally referential” paradigm is that specific and careful attention to the actual relationship between call structure and function has been sidelined by efforts to draw often metaphorical parallels with human language (e.g., Sapolsky 2017, for review see Owren and Rendall 2001).

1.2 Research Goals, and Layout of this Thesis

The vocal signals produced by vervet monkeys have long been at the center of debates surrounding the relevance of functionally referential signaling to the evolution of human language, as well as debates surrounding the utility of information-based interpretations of animal signaling as a whole (Rendall et al. 2009a; Scarantino 2010; Cheney and Seyfarth 2005; Seyfarth et al. 2010; Wheeler and Fischer 2012, 2015; Scarantino and Clay 2015). Likely as a result of this historical focus on functional reference, surprisingly little empirical work has focused on other sources of acoustic variation within these call types. This empirical gap is problematic, however, as the physical structure of vocal signals can provide important insight into both their function and the selective pressures that have shaped their evolution. Different call types are likely to be more or

less constrained in the amount and sources of variation in their acoustic structures depending on their proximate and ultimate functions. For example, calls that are used in signalling among many group members in a specific context (such as alarm calls) may have vastly differing constraints on variation than those used more commonly in dyadic, inter-individual interactions across a variety of contexts.

In the following study, I explore the acoustic variation in the vocalizations produced by vervet monkeys (*Chlorocebus pygerythrus*). In particular, I focus on two call types that have been studied intensively, but viewed almost exclusively through the lens of functional reference; grunts and terrestrial alarm calls produced in response to terrestrial predators. I set out to determine whether the acoustic structure of these signals vary in predictable ways based on the individual characteristics of the caller, including their age, sex, identity, and body weight. I also explored the effect of context on the acoustic structure of the vervet monkey's grunt vocalizations. My overarching goal was to explore sources of acoustic variation in these calls in an attempt to better understand their function, and the mechanisms by which they exert their effects on receiver behaviour.

Chapters 2, 3 and 4 are formatted as independent articles for publication. In **Chapter 2**, I explore sources of acoustic variation within the grunt vocalizations of vervet monkeys. First, I explore the degree to which grunts vary in structure in relation to a number of caller-specific characteristics, including their age, sex and individual identity. Grunts are generally produced in close range social situations, and I explore these individually-specific traits as they may play an important role in determining the outcomes of social interaction. I also revisit the contextual specificity of grunt vocalizations, taking into account a wider range of contexts than were originally considered in Cheney and Seyfarth's (1982a) first study.

In **Chapter 3**, I explore the acoustic variation in the terrestrial alarm calls of vervet monkeys, with a particular focus on age/sex differences in call structure. The terrestrial alarm calls of vervet monkeys have long been described as two distinct call types, with females and immatures producing one call type, and adult males producing a second call type (Struhsaker 1967a). Still, these differences have yet to be quantified, likely in light of the fact that from a referential standpoint, both call types functionally “stand for” the same predator classes. Still, these sex differences in the acoustic structure (form) of terrestrial alarm callers may be the product of sex specific selection for different acoustic parameters. I begin to explore this possibility by determining which acoustic parameters best distinguish between the calls of males and females. I also use a statistical clustering technique to determine whether these calls actually exist as two distinct call “types” or whether they exist as endpoints within a continuously graded system. Finally, I examine terrestrial alarm calls for evidence of individual differences in alarm structure.

Based on the results presented in chapters 2 and 3, **chapter 4** explores the degree to which sex differences in body size explain sex differences in the structure of grunts and alarm calls in adults. I do this with the goal of determining whether sex differences in either call type are simply related to body size dimorphism in adults, or whether sex differences in call structure exceed what would be expected based on simple body size dimorphism.

In the remainder of **Chapter 1**, I review and outline relevant literature surrounding acoustic variation in the vocal signals of primates. First, I outline the mechanisms by which vocal signals are produced. Following this, I will discuss the sources of individually specific variation in the acoustic structure of vocal signals that arise as a by-product of these vocal production mechanisms. Specifically, I will discuss variation in vocal structure as it relates to a caller’s body size, age, sex, individual identity, and motivation. I then focus on contextual variation in call structure, with an

emphasis on the vocal repertoire of vervet monkeys. Finally, I highlight some concerns regarding the central focus on *information transfer* in the study of animal communication systems.

1.3 Literature Review

Source Filter Theory and the Relationship Between Body Size and Acoustic Structure

A thorough understanding of the variation that characterizes the vocal signals of primates requires at least some understanding of the mechanisms by which vocalizations are produced. This is because much of the variation in the structure of vocal signals are the direct result of within and among individual differences in this vocal production system. The vocal production system of most terrestrial vertebrates consists of two, functionally independent components: the source, and the filter (Titze 1994, Fitch and Hauser 2003). Anatomically, the source is composed of the lungs, and the vocal folds in the larynx. When animals vocalize, air is expelled from the lungs, through the trachea, and into the larynx (or the analogous syrinx of birds, which is located at the base of the trachea). This causes the vocal folds to vibrate, which in turn generates sound (Fitch and Hauser 2003). In some instances, the vibrations of the vocal folds are periodic, and the frequency at which the vocal folds vibrate during phonation gives rise to a vocalization's fundamental frequency (what human perceivers interpret as pitch). This frequency is determined in part by the tension, length, mass, or thickness of the vocal folds. These vibrations can also generate harmonics, which are additional vibrations at frequencies at whole number multiples of the fundamental frequency. It is also possible for the vocal folds to vibrate in an a-periodic fashion, producing random frequencies with respect to time (called noise).

After sound is produced by the vocal folds, it passes from the larynx into the rest of the vocal tract. Sounds traveling through the supralaryngeal vocal tract are “filtered”, in that various frequencies from the source signal are amplified and/or dampened before the sound is released into the environment. The frequencies that are amplified or dampened are a function of resonances and anti-resonances of the vocal tract, which are primarily determined by the length and shape of the pharynx, and both the nasal and oral cavities. Vocal tract resonances can be modified through movements of the soft pallet, tongue, lips, and the jaw itself, in a process called “articulation” (Fitch and Hauser 2003, Briefer 2012). As the articulators in the supralaryngeal vocal tract move, the resonances and anti-resonances of the vocal tract change, modifying the characteristics of the source waveform. The size and shape of the vocal tract, in conjunction with the position of the articulators can result in concentrations of acoustic energy at distinct, and relatively narrow frequency ranges, creating spectral peaks (often called formants). Vocal tract length affects formant dispersion, with longer vocal tracts resulting in a decrease in the spacing between adjacent formants (reduced formant dispersion). By dynamically altering the position of the articulators, mammals can affect the structure of the calls they produce. For instance, some call types may be produced while callers protrude their lips, which would functionally elongate the vocal tract, thus changing its filter characteristics. Fitch (2000), Hauser et al. 1993, Hauser and Ybarra 1994). For example, rhesus macaques (*Macaca mulatta*) protrude their lips when emitting coo vocalizations but not screams (Hauser et al. 1993), and this lip protrusion lowers the vocal resonances of the calls they produce (Hauser and Ybarra 1994).

Variation in Relation to Body Size

The fundamental relationship between the physical characteristics of the vocal tract and the acoustic structure of vocal signals suggests that vocalizations are likely to vary in structure based on the physical characteristics of a caller (Fitch and Hauser 2003). As such, it is likely that in many instances, receivers may be able to extract accurate information about a caller's body size from the vocal signals they produce. Body size is likely to play an important role in many social interactions, especially when concerning competition over resources, simply because larger individuals are likely to dominate smaller individuals when contests escalate to physical aggression (Fitch and Hauser 2003, Morton 2017). Vocal signals that provide reliable cues to body size could thus act as replacements for actual physical aggression in many instances; animals will avoid engaging in aggressive interactions with individuals who produce “larger” sounds.

From a mechanistic standpoint, these cues are likely to be available in vocal signals because variation in the size of the lungs, larynx, and vocal folds affect the source waveform in predictable ways (Fitch and Hauser 2003). Because air from the lungs is required to drive oscillations of the vocal folds, lung size, which is generally correlated with body size, can have a direct impact on both the duration, and amplitude of vocal signals, making these vocal parameters potentially honest indicators of caller size within call types (Fitch and Hauser 2003). Specifically, larger animals are able to produce longer and louder vocal signals relative to smaller individuals. Additionally, based on the assumption that larger individuals should have larger vocal folds, fundamental frequency has long been thought to be an honest cue to caller size. Although this has proven to be accurate when comparing vocal signals between broad size categories as might be expected between different breeds of dog (Taylor et al. 2008), or between callers of different age or sex classes (Rendall et al. 2004; Ey et al. 2007a; Taylor and Reby 2010), fundamental frequency is generally

considered to be an unreliable cue for an animal's body size when comparing individuals within the same species, or age/sex classes. This is because the larynx is a cartilaginous structure whose growth is unconstrained by any rigid structures in the body, and can therefore grow independently of body size. This results in the potential for fundamental frequency to decouple from body size, which has been shown in a number of species (Japanese macaques, *Macaca fuscata*: Masataka, 1994; rhesus macaques, *Macaca mulatta*: Fitch, 1997; humans, *Homo sapiens*: Collins, 2000; red deer, *Cervus elaphus*: Reby and McComb, 2003; baboons, *Papio hamadryas ursinus*: Rendall et al., 2005, speckled ground squirrel, *Spermophilus suslicus*, and yellow ground squirrel *Spermophilus fulvus*: Matrosova et al. 2007). Conversely, the pharynx, and both the oral and nasal cavities of most mammals are constrained by the size of the skull and mandible, which are generally highly correlated with body size. As such, the dispersion of frequency peaks within vocal signals has proven to be a reliable cue to the size of signalers in a number of species (Japanese macaques: Fitch 1997; domestic dogs: Riede and Fitch 1999; red deer: Reby and McComb 2003; Taylor et al. 2008; pandas: Charlton et al. 2009). Still, while the relationship between body size and the resonant characteristics of the vocal tract may hold between broad size categories within species, there is evidence that the relationship between body size and the filter characteristics of the vocal tract may be less apparent when comparing individuals that are closer in overall body size (Rendall et al. 2004).

Variation Related to Age and Sex

The relationship between body size and acoustic structure means that vocal signals are likely to contain cues to a number of caller specific characteristics that are correlated with body

size. For instance, age differences in the structure of vocal signals have been found in a number of non-human primates (Ey et al. 2007a). Most evidence suggests that the acoustic structure of primate vocal signals are innate, and undergo little structural modification throughout development (Egnor and Hauser 2004). As such, where age-related variations in call structure have been found, they are generally thought to be by-products of physical growth and maturation of the various components of the vocal tract.

Research has indicated that in some instances, receiver responses are affected by these age differences in call structure. For example, the age of callers seems to affect the response of receivers to alarm calls in a number of species (California ground squirrels, *Otospermophilus beecheyi*: Hanson and Coss 2001; squirrel monkeys, *Saimiri sciurius*: McCowan et al. 2001; bobac marmot, *Marmota bobac mull*: Nesterova 1996; Belding's ground squirrels, *Urocitellus beldingi*: Robinson 1981). Responses to alarm calls produced by adults are often more pronounced than to those of juveniles (e.g., California Ground Squirrels, *Spermophilus beecheyi*: Hanson and Coss 2001; bonnet macaques, *Macaca radiata*: Ramakrishnan and Coss 2000). This is presumably due to the fact that younger individuals are more likely to misclassify predators due to a lack of experience, making their alarm calls less reliable. The opposite pattern is true in Yellow-bellied marmots (*Marmota flaviventris*), where adults increase their level of vigilance, and suppress foraging behaviour for longer periods after playbacks of juvenile alarm calls versus those of other adults, possibly because younger individuals are more vulnerable to predation (Blumstein and Daniel 2004). Still, a lack of differentiation between the alarm calls of different age classes have also been reported in some instances (e.g., Richardson's ground squirrels, *Spermophilus richardsonii*: Swan and Hare 2008). In this species, there is no evidence of age-related differences in call structure, meaning receivers are likely unable to distinguish between the calls of different

age classes. Similarly, the alarm calls of the speckled ground squirrel (*Spermophilus suslicus*) and the yellow ground squirrel (*Spermophilus fulvus*) seem to lack cues to caller age and body size (Matrosova et al. 2007). Researchers suggest that this may represent a case of vocal mimicry, in which immature animals produce adult-like calls as a means of decreasing their risk of attracting predators that specialize on younger prey, or to decrease the risk of infanticide (Matrosova et al. 2007).

Variation in body size between males and females has also been found to correlate with sex differences in the structure of vocal signals. Again, in many instances, these differences are likely the product of size-related variations in the anatomy of the vocal tract. Sex discrimination of vocalizations has also been found in a number of taxa (Eared grebes, *Podiceps nigricollis*: Nuechterlein and Buitron 1992; cotton top tamarins, *Saguinus oedipus*: Miller et al 2004; Common marmosets, *Callithrix jacchus*: Chen et al. 2009; red-chested moustached tamarins, *Saguinus labiatus*: Masataka 1987; Black-Tufted-Ear Marmosets, *Callithrix kuhlii*: Smith et al. 2009). For example, phonotaxis experiments revealed that male and female Cotton top tamarins (*Saguinus oedipus*) preferentially approach long calls produced by the opposite sex (Miller et al. 2004). Additionally, when call parameters were experimentally manipulated to exaggerate the perceived ‘male-ness’ or ‘female-ness’ of calls, subjects showed a preference for the exaggerated calls of the opposite sex, suggesting that males and females possess sensory biases that are sensitive to those sex-linked parameters (Miller et al. 2004). Humans also infer gender based on voice using a range of acoustic properties including vocal pitch and resonant characteristics of the vocal tract (Mullennix et al. 1995, Simpson 2009, Junger et al. 2013). The perception of male versus female voices in human speech has been shown to activate different regions within the brains of males

(Sokhi et al. 2005), and both men and women show a tendency towards increased attention towards speakers of the opposite sex (Junger et al. 2013).

Sexual selection and body size exaggeration

A number of studies have highlighted the fact that vocal signals thought to function in both inter- and intra-sexual selection vary in structure based on the caller's body size, fighting ability, resource holding potential, and stamina (Fischer et al. 2002, 2004, Kitchen et al. 2003, Wyman et al. 2008). In the case of male-male competition, these structural cues likely benefit both signalers and receivers by reducing the need for actual physical aggression during competitive interactions. Similarly, this acoustic variability likely provides cues from which females can select preferred mates. It seems, however, that selection for cues related to "quality" have in some instances severed the relationship between body size and call structure. For example, a number of studies have highlighted the fact that differences in the fundamental frequency of vocalizations produced by males and females exceeds what would be predicted based solely on sex differences in body size (Harris et al. 2006, Schoen 1971, Fitch and Giedd 1999, Fitch and Reby 2001, Puts et al. 2012, Delgado 2006, Rendall et al. 2005). One commonly cited explanation for this is that sexual selection has acted on the structure of male vocal signals to exaggerate body size (Fitch 1999, Taylor and Reby 2010, Puts et al. 2016). Because body size likely correlates with fighting ability or resource holding potential, it is thought that by acoustically exaggerating body size, signalers may be able to better intimidate competitors, or to influence mate choice in females (Delgado 2006). One criticism of this hypothesis, however, is the fact that fundamental frequency does not seem to track body size within age/sex classes, suggesting that selection should act on receivers to ignore this trait provided it acts as a "dishonest" cue (Taylor and Reby 2010, Fitch 1997). Still,

despite the lack of a clear correlation between body size in adult males and fundamental frequency, a number of studies have shown that at least in humans, men and women seem to take this variable into account when rating males in terms of attractiveness or relative dominance (Feinberg et al. 2005, Puts et al. 2007). Some authors have suggested that this trait may actually better represent a caller's immunocompetence, and could therefore still represent an honest cue to a caller's resource holding potential (Puts et al. 2016). Secondary growth of the vocal folds in males at puberty is mediated by testosterone (Fitch and Hauser 2003). Because the effects of testosterone are inhibited by stress-related hormones such as cortisol (Mehta and Josephs 2010), some authors have suggested that the drop in fundamental frequency in the male voice at puberty should be more pronounced in healthier males (Puts et al. 2016).

Variation Related to Individual Identity

Acoustic analysis has shown that cues to the identity of a caller are often present in the vocalizations of non-human primates. Because primates are highly social animals, their vocal signals are often exchanged between individuals who interact with one another on a repeated basis across a variety of contexts. Because social life can be characterized by highly differentiated relationships between individuals, it is likely that these indexical cues are a powerful way to influence a receiver's behavior (Tibbetts and Dale 2007). Research has shown, for example, that vocal recognition between mothers and their offspring is common among birds and mammals, particularly in species that reproduce in large social groups or colonies, where recognition of particular individuals based on location alone could be difficult (Cheney and Seyfarth 1980, Sieber 1986, Kent 1987, Balcombe 1990, Medvin et al. 1993, Maestriperi and Call 1996, Charrier et al.

2002, 2003, Sebe et al. 2007, Pitcher et al. 2010). Individual recognition of vocal signals may also allow receivers to assess the reliability of different callers (Hare and Atkins 2001, Sproul et al. 2006). For example, Hare and Atkins (2001) found that Richardson's ground squirrels (*Spermophilus richardsonii*) reduced their responsiveness to playbacks of alarm calls produced by individuals whose calls had been experimentally broadcast (using a playback protocol) in the absence of any threat. Conversely, receivers still responded to the calls of 'reliable' signalers, whose calls had been experimentally paired with the presentation of an artificial predator mount (a stuffed badger). These results suggest that one possible advantage of distinguishing between the calls of different individuals is the assessment of reliability in the signals of conspecifics. The tendency to selectively respond to signals based on an assessment of the caller's reliability has been found in a number of species (e.g., vervet monkeys: Cheney and Seyfarth 1988; bonnet macaques, *Macaca radiata*: Ramakrishnan and Coss 2000; California ground squirrels, *Spermophilus beecheyi*: Hanson and Coss 2001; marmots, *Marmota flaviventris*: Blumstein et al 2004).

Individual recognition may also allow individuals to reduce the costs associated with overt physical aggression (Barnard and Burk 1979); in species where there is the potential for repeated aggressive interactions between two individuals, receivers may come to associate the vocalizations of particular callers with a given social outcome (Owren and Rendall 2001). In this way, highly competitive individuals may be able to invoke learned fear in receivers by simply producing vocalizations that have been associated with physical aggression in the past. In these instances, receivers are likely to give way to more dominant callers without the need for costly, overt aggression (Owren and Rendall 2001). By the same mechanism, calls produced by particular individuals in association with benign interactions may also have learned mollifying effects on

receivers, thus facilitating social interactions in species where the intentions of conspecifics could be ambiguous in the absence of such cues. Caller identity also seems to play a role in territoriality, with vocal recognition of neighbors having been documented in a number of species (birds: Falls 1982, Godard 1991; bullfrog, *Rana catesbeiana*: Davis 1987). The “dear enemy” hypothesis suggests that receivers can selectively focus territorial behaviours towards strangers, as known neighbors with established territories are likely to pose less of a threat than transient individuals in search of a territory (Fisher 1954, but see Temeles 1994).

A number of primate species produce individually recognizable vocal signals collectively known as “contact-calls”, which have been hypothesized to allow social groups to remain cohesive when maintaining visual contact is difficult (Caine and Stevens 1990, Kondo and Watanabe 2009). These calls may allow individuals to both monitor, and attract the attention of particular group members, facilitating cohesion and the coordination of behaviours between closely bonded individuals. In many instances, primates have shown differential responses to these vocalizations based on their relationship with a particular caller. For example, Ramos-Fernandez (2005) found that spider monkeys (*Ateles geoffroyi*) were more likely to approach playback speakers broadcasting whinny vocalizations if the calls were originally produced by a close associate of the receiver. Similarly, Rendall et al. (1996) found that caller identity influenced how rhesus macaques respond to ‘coo’ vocalizations, with females responding more strongly to calls produced by their own kin. Screams produced during antagonistic interactions have also been found to contain individually specific cues (e.g., rhesus monkeys: Gouzoules et al 1984, Fugate et al. 2008; chimpanzee: Slocombe and Zuberbühler 2005; vervet monkeys: Cheney and Seyfarth 1980), and may allow callers to recruit kin or coalition partners for support, in a number of species (Cheney 1977; Gouzoules et al. 1984; Bernstein and Ehardt 1985; Gouzoules et al. 1985, 1986).

Experimental evidence from vervet monkeys suggests that in some instances, the ability to recognize individuals based on their vocalizations may even extend to individuals outside the receiver's social group (Cheney and Seyfarth 1982b). Experimental playbacks of artificially constructed vocal sequences representing rank reversals in baboons suggest that primates not only attend to acoustic cues to identity and kin-group membership, but that "eavesdropping" individuals also monitor the social interactions of specific individuals based on these vocal cues (e.g., Cheney et al. 1995, Bergman et al. 2003).

Mechanistically, it is likely that cues to caller identity are the result of idiosyncrasies in the structure of the vocal tract of different individuals, which affect the physical properties of their vocal signals in individually-specific ways. A number of studies have highlighted the fact that within a species' vocal repertoire, some call types provide more reliable identity cues than others (Charrier et al. 2001, Rendall et al. 2009, Bouchet et al. 2012). A number of hypotheses have been proposed as to why some signals may be under stronger selection pressure to contain cues to caller identity. One hypothesis centers around the idea that long distance vocalizations should be under stronger selection to contain individually specific cues than close range vocalizations (Marler 1967, Bouchet et al. 2012). This is because when callers and receivers are separated, vocal signals may be the only way by which signalers can locate or communicate with specific individuals. Conversely, animals are presumably able to use visual cues to assess caller identity when in close proximity, meaning there may be less selection pressure for clear individual differences in call structure in calls used at close range. Of course, call function is also likely to play a role in the importance of individual variability. As such, some authors have hypothesized that calls that function primarily in social situations should contain clearer cues to caller identity than calls emitted in non-social contexts (e.g., alarm calls) (Charrier et al. 2001). This hypothesis has been

supported by a number of studies that examined the degree of individual distinctiveness between call types within the same species; calls whose functions were related to social interaction generally contained clearer cues to call identity than calls with less direct social function (Charrier et al. 2001).

Variation related to Motivation

The vocal signals produced by non-human animals have often been viewed as reflections of a caller's internal state, acting as graded expressions of a signaler's current motivation, emotional state, and/or their level of arousal (Rowell and Hinde 1962; Lancaster 1975; Premack 1975; Smith 1977). The link between motivational state and acoustic structure was formalized by Morton (1977), who noted that the vocalizations of a number of bird and mammal species varied in predictable ways based on the presumed motivational state of the caller (e.g., hunger, sexual arousal, active aggression, retreat). Specifically, signals emitted in agonistic situations tend to be relatively harsh, long, and are characterized by an energy distribution predominantly in the lower frequencies. Conversely, when signalers are in non-aggressive or fearful situations, signals are often less noisy, shorter, and have the majority of their energy focused at higher frequencies. These generalizations were termed "motivational structural rules" (Morton 1977), and were thought to be the by-product of selection for signal structure which could act on the pre-existing sensory biases of receivers. For instance, the acoustic qualities associated with "hostile" contexts were thought to emulate the sounds produced by larger animals, therefore increasing the perceived size of callers in contexts where larger body size confers some benefit. Conversely, sounds emitted in fearful or friendly contexts were thought to mimic the sounds produced by infants to some extent,

thereby having an appeasing effect on receivers. A number of studies have tested and confirmed the predictions of this model in several species (e.g., African wild dog, *Lycaon pictus*: Robbins and McCreery 2003; coati, *Nasua narica*: Compton et al. 2001; greater tube-nosed bats, *Murina leucogaster*: Lin et al. 2015; white-faced capuchins, *Cebus capucinus*: Gros-Louis et al., 2008, humpback whales, *Megaptera novaeangliae*: Dunlop 2017).

From a functional point of view, callers are likely able to influence a receiver's behaviour by providing cues to their internal state simply because there is likely a correlation between internal state and a caller's upcoming behaviour. As such, signalers can, for instance, avoid costly physical aggression by simply providing cues to their hostile or aggressive motivation; if a receiver is unlikely to win a physical fight with the signaler, it is likely they will back-down at the threat of aggression. Importantly, cues to motivation are also extremely important in linguistic communication, and have a significant effect on the perceived meaning of a given phrase. For instance, the words "you're pregnant?" versus "you're pregnant!" can have two very different meanings for the receiver based on motivational cues available in the speaker's inflection.

The source-filter model also allows researchers to generate predictions about the impacts of a caller's affective state on the acoustic structure of the vocal signals it produces (Taylor and Reby 2010, Briefer 2012). Change in internal state may be reflected in vocalizations due in part to involuntary movements, and changes in the tension of muscles that are involved in vocal production. Changes in emotional state may also affect vocal production through changes in a caller's rate of salivation, and respiration (Scherer 2003). For example, variation in respiration may result in changes in the amplitude, duration, or rate of a signaler's vocalizations. Increased pressure from the lungs during phonation can increase the fundamental frequency of a vocalization, while increased tension of the muscles involved in respiration can result in longer

call duration and amplitude (Briefer 2012). Changes in salivation rates can have an effect on the resonance of the vocal tract, resulting in changes in formant frequencies (Scherer 2003; Scherer 1986; Briefer 2012). Vocal cues to arousal include changes in fundamental frequency (F0), call duration, call rate, and call amplitude. Vocal cues such as higher amplitude, increased F0, increased range in F0, variability in F0, and shorter inter-vocalization intervals (increased call rate) are associated with high emotional arousal (Scherer 1986; Murray and Arnott 1993; Bachorowski and Owren 1995a,b; Banse and Scherer 1996). Vocal cues associated with valence conversely are often tied to the filter, including variables such as energy distribution, and formant frequencies (Banse and Scherer 1996; Laukkanen et al. 1997; Pollermann and Archinard 2002; Waaramaa et al. 2006, 2010). For example, when compared to vocalizations associated with negative emotions, vocalizations associated with positive emotions tend to be less noisy, show narrower frequency ranges, and earlier positions of peak frequency (Hammerschmidt and Jürgens 2007; Goudbeek and Scherer, 2010).

Variation related to Context

A tremendous amount of research into the vocal signals of non-human animals has focused on what are termed “functionally referential” signals (reviewed in Marler et al. 1992, Evans 1997, Blumstein 1999, Zuberbühler 2000, Kirchhof and Hammerschmidt 2006, Wheeler and Fischer 2012, Cäsar and Zuberbühler 2012, Gill et al. 2013, Townsend and Manser 2013, Zuberbühler and Neumann 2017). These signals are generally thought to provide receivers with information regarding the contexts surrounding their production, due to the fact that they are consistently produced in association with a particular event (Macedonia and Evans 1993). Interest in the

possible referential qualities of animal vocal signals became a central topic in the study of primate communication following a series of studies on the vocal signaling behaviour of vervet monkeys (then classified as *Cercopithecus aethiops*, and since re classified as *Chlorocebus pygerythrus*) in Amboseli park, Kenya (Seyfarth et al. 1980a,b; Cheney and Seyfarth 1980, 1981, 1982a,b, 1985; Seyfarth and Cheney 1986, 1990). Probably the most well-known of these studies is the now classic work examining the production and responses to vervet alarm calls (Seyfarth et al. 1980a,b). This study was based on previous descriptive work, which suggested that vervet monkeys appeared to possess multiple, structurally discrete classes of alarm calls, which were each respectively emitted in response to the detection of different predator classes (Struhsaker 1967a). In a series of playback experiments designed to test the responses of vervets to these calls empirically, Seyfarth et al. (1980a,b) found that when alarm calls were played back to receivers in the absence of an actual predator, their responses were functionally adaptive given the class of predator that originally elicited the call. Specifically, in response to a leopard alarm call, vervet would climb into trees where they were presumably harder to catch by large mammalian predators. Calls produced in response to snakes would elicit bipedal scanning behaviour towards the ground, and calls to aerial predators would cause receivers to leave trees, and take shelter under shrubs (Seyfarth et al. 1980a,b).

Because receivers responded to the playbacks as if they were aware of the specific class of predator associated with each call type, it was suggested that the calls contain meaningful information about not only the presence of a predator, but about the specific class of predator as well. This interpretation differed significantly from how animal vocalizations had been seen traditionally, in that the alarm calls seemed to represent, or ‘reference’ a specific event or object in the environment, and were not simply an expression of a signaler’s internal state, which was

only interpretable alongside additional contextual information surrounding the call's production. The alarm calls were suggested to act symbolically, as they appeared to be arbitrary with respect to their referents in much the same way that the structure of human words are arbitrary with respect to their meaning; in much the same way that there is nothing inherent about the sound of the word "cookie" that makes it well adapted to stand for a small, round, delicious comestible, there was nothing about the sound of an "eagle rraup" that made it more or less appropriate to be associated with an eagle than it was, for instance, to be associated with a snake or a leopard. Soon after their work with vervet alarm calls, Cheney and Seyfarth (1982a, Seyfarth and Cheney 1984) found that grunt vocalizations produced by vervet monkeys varied in structure based on their context of use, and playback experiments suggested that each grunt variant produced different responses in receivers, suggesting that the calls may also function in a "referential", or "symbolic" manner.

Functionally referential signalling

The interpretation that animal vocal signals contain representational information has been met with a great deal of contention. Early criticism in particular took issue with the then-unsubstantiated conceptual parallels that were drawn in the absence of empirical evidence between vocal communication in nonhuman animals and human language. Currently, there is a more robust consensus that animal vocal signals should not be considered as homologous, or even analogous to human language (Seyfarth and Cheney 1997, Cheney and Seyfarth 1998, Hauser et al. 2002, Seyfarth and Cheney 2003, but see Scarantino and Clay 2015, Scarantino 2013, Townsend and Manser, 2013). This view largely centers on the fact that, because there is little evidence that most animals possess a theory of mind (Tomasello and Call 1997; Penn and Povinelli 2007, Martin and Santos 2016, but see: Flombaum and Santos 2005, Hare et al. 2000, 2006, Melis et al. 2006, Santos

et al. 2006, Hare et al. 2001, MacLean and Hare 2012), it is likely that signalers are not motivated to call with the intention of informing naïve receivers (Cheney and Seyfarth 1990a,b; Rendall et al. 2009a; Scarantino 2010). As the intention to inform is considered a central motivation in the use of human language, a lack of theory of mind precludes animals from possessing language on a cognitive level (Cheney and Seyfarth 2005).

One way of addressing the differences between human words and animal vocal signals has been a change in terminology; signals that seem to transmit specific information to receivers about events or objects in the external world began to be referred to as being “functionally referential” (Marler et al. 1992, Macedonia and Evans 1993), with the term “functionally” acting as an admission that although the calls seemed to carry some form of information, the nature of this information, or the mechanism by which this information was transmitted to receivers was not understood. A signal is considered to be functionally referential if it meets two basic criteria: stimulus specificity (aka: context specificity), and context independence (Macedonia and Evans 1993). The first criterion, ‘stimulus specificity’, is met if a given signal is produced in only a narrow range of contexts. Regardless of how broad or narrow this range of contexts, ‘stimulus specificity’ suggests that a signal is not produced at substantial rates in contexts that fall outside the appropriate category. The criterion of stimulus specificity allows receivers to ascertain the eliciting context of a signal due to the predictable relationship between a particular signal and a given stimulus. The criterion of context independence is met if receivers respond to a signal as if they themselves have experienced, or are aware of, the eliciting stimulus, regardless of whether or not they perceive the stimulus themselves. This is exemplified by descriptions of functionally referential alarm calls, where receivers respond to, for example, aerial alarm calls by behaving in a manner that decreases their chances of being taken by a predator that hunts from above, as

opposed to from the ground (Seyfarth et al. 1980a,b; Macedonia 1990). To date, the clearest examples of functional reference generally come from predator alarm calls (e.g., vervet monkeys; Seyfarth et al. 1980a,b; Diana monkeys, *Cercopithecus diana*: Zuberbühler et al. 1997; ringtailed lemurs, *Lemur catta*: Macedonia 1990, Pereira and Macedonia 1991; Campbell's monkeys, *Cercopithecus campbelli*: Zuberbühler 2002; black-and-white colobus monkeys, *Colobus guereza*: Schel et al. 2009, 2010; Gunnison's prairie dogs, *Cynomys gunnisoni*: Kiriazis and Slobodchikoff 2006; meerkats, *Suricata suricatta*: Manser 2001; tufted capuchin monkeys, *Cebus apella nigrinus*: Wheeler 2010). Still, functionally referential calls have been identified in food calls (rhesus macaques, *Macaca mulatta*: Hauser and Marler 1993, Hauser 1998; Geoffroy's marmosets, *Callithrix geoffroyi*: Kitzmann and Caine 2009; chimpanzees, *Pan troglodytes*: Slocombe and Zuberbühler 2005, 2006; Bonobos, *Pan paniscus*; Clay and Zuberbühler 2009, 2011) and in calls emitted in social contexts as well (vervet monkeys: Cheney and Seyfarth 1982a; rhesus macaques: Gouzoules et al. 1984; chacma baboons, *Papio cynocephalus ursinus*, Owren et al. 1997, Rendall et al. 1999; dogs, *Canis familiaris*: Faragó et al. 2010).

Information and Influence

To a number of authors, the concept of information transfer is central to the study of vocal communication (Seyfarth et al. 2010, Font and Carazo 2010). Advocates of the information-centered perspective suggest that the existence of predictable relationships between the structure of a signal (vocal or otherwise) and some relevant feature of the signaler or the environment (e.g., caller identity, age, sex, competitive ability, motivational state, context) allows receivers to make inferences about these features upon perceiving a signal. In this sense, signals are seen as providing

“information” to signalers, in that the signal effectively reduced the receiver’s uncertainty about some aspect of the world around them (Seyfarth et al. 2010). Still, a number of researchers criticize information-based interpretations of animal signaling on several levels. First, some authors highlight the fact that the concept of “information transfer” evokes a strong sense of cooperation between signalers and receivers, where in many instances, it is likely that their fitness goals do not overlap (Dawkins and Krebs 1978, Krebs and Dawkins 1984). Additionally, information-based interpretations pre-suppose that vocal signals have evolved to transfer information between individuals, despite the fact that signal function can often be achieved without the need for information transfer to have actually occurred (Morton 2017). For instance, Owren and Rendall (1997) point out that information centered views of animal signaling ignore the fact that signals can also affect receiver behaviour directly, based on the direct effects of acoustic energy on the receiver. This idea centers around the fact that sound on its own is able to evoke emotional or physiological responses in receivers. The acoustic startle reflex (Yeomans and Frankland 1995) for instance highlights the fact that sound itself can directly affect a receiver’s state of arousal. Owren and Rendall (1997, 2001) highlight the fact that alarm calls are generally characterized by fast onsets, and often consist of multiple pulses. This specific acoustic structure may be tuned to impact the central nervous system of receivers directly, causing receivers to reorient their attention towards the source of the call, while at the same time elevating their level of arousal in preparation for an escape response (Owren and Rendall 2001). Similarly, the harsh screams produced by individuals being physically aggressed by conspecifics are loud, abrupt, high intensity, and noxious. These acoustic characteristics are thought to make victims un-desirable targets for aggressors simply due to the overall aversive qualities of the calls (Rendall et al. 2009b). This interpretation differs substantially from information-based accounts which, as discussed above,

might suggest that the function of screams is to convey information about a caller's identity and level of distress, thereby functioning to elicit aid (e.g., Slocombe and Zuberbühler 2007). While it is of course possible that close associates of victims may be inclined to come to a caller's aid after hearing their screams, information-based definitions of signals ignore the jarring effects of the scream on the attacker altogether. Similar non-information-based mechanisms have also been proposed for signals produced in affiliative contexts. For example, the purr-like vocalizations produced by racoons (*Procyon lotor*) are generally produced in association with social grooming, copulation, and nursing. Purring generates rhythmic, physical pulses of the caller's body, and it has been suggested that the calming effects of purrs on receivers are likely related to tactile stimulation (Sieber 1984). Again, information-centered views ignore the possible physiological effects of these vocalizations and their concomitant vibrations on receiver behaviour.

The complex and specific behavioural responses to vocal signals can also be explained by evoking the concept of classical conditioning, whereby the consistent pairing of a vocal signal with a particular social or ecological outcome creates an association between the signal and a particular affective state in receivers (Owren and Rendall 2001). For instance, vocalizations such as grunts may facilitate social interactions among individuals not because they convey specific information about a caller's intentions or their current behaviour, but simply because grunts are generally associated with a particular social outcome. If grunts are generally associated with peaceful approaches from callers, then they will likely elicit positive emotional states in receivers due to simple associative learning. Similarly, calls generally associated with aggressive encounters may elicit appropriate avoidance behaviours in receivers based on their previous experience with callers producing those signals. This is conceptually not unlike the idea of an 'emotional contagion', whereby a vocal signal associated with a particular emotional state has the effect of inducing that

emotion in receivers. For example, Provine (1992) found that experimental playbacks of laughter elicited smiles and laughter from human subjects in the absence of other stimuli, suggesting that laughter itself has the ability to elicit common emotional states in receivers.

1.4 Brief outline of subsequent chapters

In the following three chapters, I explore sources of variation in the acoustic structure of grunts and terrestrial alarm calls produced by vervet monkeys. In **Chapter 2**, I explore sources of acoustic variation in grunts. I investigate whether grunts vary in structure based on a number of sender-specific attributes, including caller age, sex, and identity. Additionally, I re-assess the degree to which vervet grunts vary in structure based on the contexts in which they are produced. In **Chapter 3**, I quantify sex-differences in the acoustic structure of the vervet monkey's terrestrial alarm calls. Specifically, my goal is to determine what variables best distinguish between the calls of males and females, and whether the terrestrial alarm calls of vervet monkeys exist as two distinct call "types", or whether they are characterized by intergradation in acoustic structure. In **Chapter 4**, I investigate whether sex differences in the structure of alarm calls and grunts can be explained by sex differences in caller body size, or whether sex differences in call structure are more pronounced than what would be expected based on body size dimorphism alone.

Chapter 2 Sources of acoustic variation in the grunt vocalizations of wild vervet monkeys

2.1 Introduction

As a highly social Order, Primates produce a number of vocalizations during close-range social interactions among conspecifics (Cheney and Seyfarth 2018). These intra-group vocalizations are often lower in amplitude relative to other call types (e.g., long calls), and tend to be produced across a wide variety of contexts such as group movement, social approaches (Rendall et al. 1999), feeding, separation from group members, interacting with neighbours (Clarke et al. 2015), rest, travel (Crockford et al. 2017), and encountering predators (Clarke et al. 2015, Crockford et al. 2017). From a functional perspective, close-range calls are thought to facilitate social interactions among group members by reducing ambiguity surrounding the upcoming behaviors of conspecifics, or by helping individuals coordinate behaviors in instances where the fitness goals of group mates overlap (Cheney and Seyfarth 2018). Researchers have found many instances where the acoustic structure of close range vocalizations differ reliably between callers based on their age, sex, or individual identity (rhesus macaques, *Macaca mulatta*: Rendall et al. 1998, Hauser 1991, Hammerschmidt et al. 2000; red-bellied lemurs, *Eulemur rubriventer*: Gamba et al. 2012; Mountain gorillas, *Gorilla gorilla beringei*: Seyfarth et al. 1994; chacma baboons, *Papio ursinus*: Rendall et al. 2003, 2005; hamadryas baboons, *Papio hamadryas*: Pfefferle and Fischer 2006; Sooty Mangabeys, *Cercocebus atys*: Range and Fischer 2004; For review: Ey et al. 2007). In such instances, this variation may allow receivers to respond to vocal signals with appropriate behaviors, based on their relationship with a particular caller, or based on age or sex

specific differences in social relationships, or priority of access (Kappeler 1990, Sterck et al. 1997, Moore et al. 2003).

The mechanisms by which this individually specific variation is encoded into vocal signals are reasonably well understood; the acoustic properties of vocal signals are influenced by variations in the structure of the vocal tract (Fitch and Hauser 2003, Taylor and Reby 2010). As such, inter-individual variation in the vocal tracts of different callers will influence the acoustic structure of the vocal signals they produce. Cues to age and sex are also likely present in vocal signals because of the broad differences in the size and structure of the vocal tract which generally characterize different age/sex classes (Ey et al. 2007a). As the nature of primate social interactions can be highly dependent on the relationship between specific interactants, these individual cues may be highly relevant to receivers, especially during close range interactions.

In addition to this sender-specific information, numerous studies have identified instances where, upon perceiving a vocalization, receivers seem to respond as if they are somehow aware of the contextual circumstances surrounding the call's production. Mechanistically, it is believed that receivers may be able to infer the context in which a vocalization is emitted, so long as the signal exhibits a high degree of context specificity, meaning there is a predictable relationship between a particular signal type and a given context (Macedonia and Evans 1993, Evans 1997, Seyfarth et al. 2010). Interest in these “functionally referential” (Macedonia and Evans 1993) signals began with studies that showed vervet monkeys (*Chlorocebus pygerythrus*) produce three acoustically distinct alarm vocalizations, and that receivers respond to each call type with what appeared to be functionally adaptive escape behaviors given the type of predator that elicited the signal (Seyfarth et al. 1980a,b).

Although much of the evidence for functionally referential signals comes from alarm calls (for a review; Townsend and Manser 2013), evidence for context specific calls in close range social communication has been found in some species as well (e.g., chimpanzees, *Pan troglodytes schweinfurthii*: Crockford et al. 2018, gibbons, *Hylobates lar*: Clark et al. 2015, vervet monkeys, *Chlorocebus pygerythrus*: Cheney and Seyfarth 1982a, Seyfarth and Cheney 1984, baboons, *Papio cynocephalus ursinus*: Rendall et al. 1999). Chimpanzees for example, produce quiet ‘hoo’ vocalizations in at least three different contexts: when resting, when traveling, and upon encountering potentially dangerous hidden threats (e.g., snakes). Calls given in each context can be distinguished based on their acoustic properties, and receivers respond with contextually appropriate behaviors suggesting they can extract contextual information from these signals (Crockford et al. 2015, Crockford et al. 2018). For instance, ‘travel hoos’ seem to initiate approach and traveling behavior between closely associated individuals, while ‘rest hoos’ elicit vocal responses, but not traveling. Finally, ‘alert hoos’ elicit cautious approach behaviors from receivers (Gruber and Zuberbühler 2013, Crockford et al. 2015, Crockford et al. 2017).

Vervet Grunts

Throughout the day, vervet monkeys produce pulsatile, low pitched grunt vocalizations (Cheney and Seyfarth 1982a). Cheney and Seyfarth (1982a, Seyfarth and Cheney 1984) reported that vervet grunts were generally produced in one of four different contexts: 1) by subordinate individuals when being approached by a dominant individual, 2) by dominant individuals when being approached by subordinates, 3) by individuals when they detect another social group within their proximity, and 4) by individuals as they, or members of their group, begin moving into an

open area. Play-back experiments showed that receivers respond differently to grunts based on which of these four contexts they were originally produced (Cheney and Seyfarth 1982a). Responses differed in terms of the movements of receivers relative to the position of the playback speaker (the simulated caller), the direction the receivers' gaze after hearing the call, the longevity of a given response, and the latency between the playback and these responses. Though subtle, these differences in response were taken to suggest that vervets are able to distinguish between grunts given in different contexts based on their acoustic properties alone. Acoustic analysis revealed that grunts contained spectral peaks that varied in frequency position over the course of the call based on the context in which they were emitted (Seyfarth and Cheney 1984). This led to a linguistically-inspired interpretation of these signals, whereby grunts were seen as possessing some form of rudimentary referential qualities. The little work that has been done on grunts since the work of Cheney and Seyfarth (1982a, Seyfarth and Cheney 1984) looked specifically at grunts produced by adult males during dyadic encounters, and suggested that males may produce grunts as a signal of benign intent, or to recruit close associates as a means of reducing predation risk (Mercier et al. 2017).

Despite this work on vervet grunts, a number of issues remain unresolved. For instance, possibly as a result of the extensive interest and research effort surrounding functional reference in animal signaling, it has yet to be determined what individually-specific information may be available to receivers upon hearing the grunts of conspecifics. Because vervets emit grunts in a number of social situations - and produce them at relatively high rates throughout the day (personal observation) - it seems that one possible function of the call may simply be to maintain vocal contact with other group members. As a highly social species where the relationships between different individuals vary a great deal (Cheney and Seyfarth 1985a, 1986, Young et al. 2017), it is

likely that vervets would benefit from not only monitoring the location of their social group as a whole, but also from monitoring the whereabouts of particular individuals. Vocalizations that are highly important in social situations are predicted to be more individually distinctive compared to other call types, specifically because identity is thought to play a deterministic role in many social interactions (Charrier et al. 2001). Similarly, as the social dynamics of primates are likely to vary both within and between sexes and age groups, it is likely that vervets could benefit from both sending and receiving this information. Even from a non-adaptationist perspective, it is possible that cues to age, sex, and identity are present within the grunt vocalizations of vervets as the simple by-products of inter-individual (and inter-sex, inter-age) variations in their vocal production systems.

Previous studies that focused on the referential qualities of grunts only analyzed a subset of calls, omitting those produced during aggressive interactions, and those produced in bouts of more than one call (Cheney and Seyfarth 1982a, Seyfarth and Cheney 1984). Under natural circumstances, however, vervet monkeys are also exposed to these grunts and, as such, may be exposed to a wider range of acoustic variation in grunt structure than was considered in those early studies. Research has shown that the assessment of context specificity can be affected by the number of contexts considered in an analysis. For example, statistical classification of baboon grunts emitted in different contexts was subject to more errors when a larger number of contexts were considered (Meise et al. 2011). Specifically, these authors re-examine Owren et al. (1997), who found consistent differences in the acoustic structure of grunts produced in two distinct contexts; grunts produced as females approached mothers and their infants, and grunts produced as individuals began to move into open areas within their home-range. While this was taken to suggest that grunts encoded specific information regarding their contexts of production, Meise et

al. (2011) found that when grunts emitted in other contexts were considered in the analysis (e.g., grunts emitted while foraging or grooming), call contexts were frequently misclassified. This suggests that contextual information in these signals was actually relatively ambiguous, and that when grunts emitted across a wider range of contexts were considered, it may be difficult for receivers to infer context based on the acoustic structure of grunts alone. Similarly, the assessment of context specificity for vervet alarm calls varied when alarms were analyzed alongside acoustically similar calls produced in aggressive contexts (Price et al. 2015); when alarm calls were analysed alongside qualitatively similar calls produced in non-predatory contexts, statistical assessment of context specificity decreased. Similarly, it is possible that the association between context and acoustic structure of vervet grunts is less accurate and consistent than earlier accounts suggest, as earlier studies only analysed grunts produced in a limited number of contexts.

In this chapter, I investigate the sources of acoustic variation in the vervet monkey's grunt vocalization. First, I determine whether the acoustic structure of grunts is dependent on a caller's age, sex, or identity. Next, I examine the degree to which calls produced across different contexts vary consistently in acoustic structure. My aim is to determine whether vervet grunts contain reliable cues to caller characteristics or context.

As vervet monkeys are characterized by sexual dimorphism at adulthood (Turner et al. 1997), it is likely that sex differences in body size are associated with corresponding sex differences in vocal tract anatomy. Consequently, I expect to find sex differences in the structure of adult grunts. Similarly, as vervets increase in size throughout growth and development, I predict that grunts vary in structure as a function of caller age. I predict that variables relating to the overall frequency content of signals, as well as those relating to call (or element) duration will best distinguish between different age/sex classes, as these variables are those that are predicted to vary

as a consequence of caller body size (Ey et al. 2007, Fitch and Hauser 2003). Specifically, grunts should contain more energy at lower frequencies, and should be longer in duration in larger individuals relative to smaller individuals (i.e., in males relative to females, and in older individuals relative to younger individuals). Additionally, it is likely that the vocal tracts of different individuals are characterized by idiosyncratic variations in structure. This, in conjunction with the fact that grunts are often produced in social contexts where individual identity is likely to play an important role, suggests that there might be selection for individually-specific cues in the structure of these calls.

Based on the findings of Cheney and Seyfarth (1982a, Seyfarth and Cheney 1984), I would predict that grunts will vary in structure based on their context of production. Still, because my analyses include calls produced in both aggressive contexts, as well as grunts produced in bouts of multiple calls (which were not taken into account by those original studies; Cheney and Seyfarth 1982a, Seyfarth and Cheney 1984), it is possible that my assessment of context specificity will reveal a less consistent link between context and acoustic structure, as has been demonstrated by other studies which included a larger sample of contexts into their analyses (e.g., Meise et al. 2011, Price et al. 2015).

2.2 Methods

Study Species

Vervet monkeys are semi-terrestrial, old world monkeys, which range through-out sub Saharan Africa. The vervet monkey's northern distribution stretches from Senegal to Ethiopia, and they are found as far south as South Africa (Struhsaker 1967b). After humans and baboons, vervet monkeys are the most widely distributed primate in Africa (Struhsaker 1967c; Wolfheim 1983).

Vervets are found primarily in riparian woodland (Struhsaker 1967b), their presence being limited primarily by their access to water, and appropriate sleeping sites (Wrangham 1981). Their home range sizes vary from approximately 0.06 km² to 1.78 km² (Harrison, 1983; Willems and Hill, 2009). They defend their home ranges through the use of intergroup aggression, especially during periods of food scarcity (Chapman and Fedigan 1984). Vervets travel between 135 m to 2251 m a day, varying primarily based on the size of a troop's home range, and the availability of water and food resources (Struhsaker 1967c; Wrangham 1981). They are considered dietary generalists, consuming leaves, berries, flowers, seeds, nuts, shoots, fungi, birds, bird eggs, invertebrates and vertebrates (Struhsaker 1967c; Chapman et al. 1988; Wrangham and Waterman 1981). Vervet monkey social organization is characterized by multi-male, multi-female groups, and female philopatry. Males emigrate multiple times throughout their lives; first, as subadults from their natal group, and again as reproductively mature adults (Cheney and Seyfarth 1983). Social group size averages 25 individuals (Fedigan and Fedigan 1988). Female dominance hierarchies are linear and relatively stable, with rank being inherited from the mother (Struhsaker 1967b). Male hierarchies vary over time (Cheney and Seyfarth 1989).

Vervet monkeys are sexually dimorphic at adulthood, with females averaging 3.3kg, and males 5.5kg at my study site (Pasternak et al. 2013). Males and females begin to diverge in size after approximately 15-18 months of age (Turner et al. 1997). By 38 - 40 months of age, females have generally reached their maximum body size. Conversely, males continue to grow after this age, and they experience faster growth rates than females from 20 to 48 months of age (Turner et al. 1997). Females generally reach reproductive maturity at 30-36 months of age (Turner et al. 1997). Males generally reach reproductive maturity around 5 years of age (Horrocks 1986).

Study Site and Population

Data for this study were collected at the Samara Game Reserve, Eastern Cape Province, South Africa (32° 22'S, 24°52'E), from May 2016 to May 2017. Animals at the site are fully habituated to the presence of observers. The population at Samara game reserve has been a part of an ongoing study since 2008 by the Barrett-Henzi lab (University of Lethbridge, Psychology). Research at the site focuses on a broad array of topics relating to social dynamics, group coordination, and phenotypic plasticity. Samara is located approximately 260km north of Port Elizabeth, on the southern portion of the Sneeuwberg mountain range. It has a semi-arid climate, and is comprised of 27 000 ha of mountains and nama-karoo grassland (Pasternak et al., 2013). Data were collected from three habituated groups of vervet monkeys (Picnic Troop [PT], River Bend Mob [RBM], and Riverside Troop [RST]). Two of the groups have been habituated to researchers since 2008, and the third group has been habituated since 2012. All three study groups live along the Milk River in a semi-arid riverine woodland (Pasternak et al., 2013). Records of all births, deaths, and migration events have been recorded since the onset of data collection. Dominance interactions, including displacements, supplants, and physical aggression, are recorded on an all observed occurrence basis at the field site in order to determine dominance hierarchies. Dominance hierarchies were calculated for each of the three study groups using 'DomiCalc' (Schmid and Vries 2013) to calculate standardized David's scores in three-month blocks.

For this study, animals were classified as one of four age classes based on their year of birth; yearlings (one to two years of age- Born in 2015), juveniles (two to three years of age - Born in 2014), subadults (three years to reproductive maturity – Born in 2013), or Adult (reproductive maturity). There was one individual who was born into the study population in 2012, making him one year older than the rest of the subadults in this study. Still, this individual had yet to reach

reproductive maturity, and was included in the subadult age category for this study. The size and composition of each of the study groups varied over the course of the study, and are presented in Table 2.1.

Table 2.1: Group composition over the course of data collection. Each cell represents the range of counts for each age/sex class in each group. Variation in values is the result of Births, Deaths, Immigrations and Emigrations.

Social Group	Picnic Troop		River Bend Mob		Riverside Troop	
Sex	Male	Female	Male	Female	Male	Female
Yearling	2	0	4	3-4	2-3	3
Juvenile	5	5	5	6	4-5	4
Sub-adult	8	3-4	4	5	4-5	6
Adult	6-9	8-9	4-5	6-10	8-10	9-13

Vocal data collection in the field

I recorded vocalizations in the field using a Marantz PMD661 digital field recorder, and Sennheiser ME67 directional condenser microphone. Recordings were made at a sampling rate of 48kHz, and a bit rate of 1536kbps. The microphone was equipped with a blimp windshield and shock mounting system with a pistol grip (Sennheiser MZS20-1 Combo Mount/Grip/Stand) to reduce wind and handling noise. Each day in the field, I followed one of the three study groups continuously for a ten-hour period. In the winter, this ten-hour period coincided with daylight hours, meaning the subjects were followed from when they first left their sleeping trees until they returned to their sleeping trees at the end of the day. In the summer months, daily group follows were balanced between a) days that started at sunrise and ended ten hours later, and b) days that ended at sundown, having started 10 hours earlier. I made audio recordings and collected data regarding the production of grunts both opportunistically and during focal animal samples

(Altmann 1974). The distance between the microphone and the caller was generally between 3m and 7m when subjects were on the ground, or 3m to 10m when the subjects were in the trees.

When a vocalization was recorded (either during a focal sample, or while recording opportunistically), I would record the identity of the caller, and its behavior at the time of calling. Behaviors scored during this study are listed and described in the attached ethogram (see Table A.1, Appendix to chapter 2). I collected Ad libitum data on potentially relevant socioecological contexts surrounding the production of each call, including the presence of another social group, and notes on the general behavior of the group as a whole (see Table A.1, Appendix to chapter 2). I also collected data pertaining to the proximity of other individuals relative to the caller whenever a grunt was recorded. I collected these data as previous work on this species had suggested that grunts were often elicited in response to the proximity of both dominant and subordinate individuals (Cheney and Seyfarth 1982a). I attempted to identify as many individuals surrounding a calling animal as possible; the distance at which I could confidently identify monkeys surrounding a caller varied greatly based on many factors, including the density of foliage in the area, the height of the caller in a tree/shrub, or the density/activity of neighbors within sight. When an animal called, I would start by identifying any monkeys within body contact of the caller, followed by monkeys <1m, 1m to 2.9m, 3m to 4.9m, 5m to 9.9m, 10m to 14.9m, 15m to 19.9m, and then >20m. I would stop identifying monkeys when I felt I could no longer realistically see all the potential individuals within a given distance of the caller. I am confident that I could always identify individuals within a 5m radius of the caller. In some instances, vocalizations were given while a caller was clearly attentive of another individual (e.g., calls given by a monkey being aggressed, calls given by monkeys while being displaced by dominant individuals, or when a monkey was simply staring at another individual). I considered a monkey to be attentive of another

individual if it was clearly looking at that individual while calling (facing that individual with eyes open). In these instances, I recorded the identity of the monkey that was the subject of the caller's gaze (hereafter 'target'). Target individuals were at times further from the caller than 5m. In the case of calls given in association with social interactions (e.g., grooming, aggression, displacements), I would record both the nature of the interaction, as well as the ID of all interactants involved. All data were recorded using handheld Trimble Nomad, or Trimble Juno GPS-enabled data-loggers, equipped with Pendragon Forms Software (v 5.1).

2.3 Data Analysis

Assigning Contexts to calls

Based on my observations in the field, I distinguished between 5 contexts that were associated with the production of grunts: 1) aggression, 2) inter-troop encounters (ITE), 3) presence of a dominant individual (Dominant), 4) presence of a subordinate individual (Subordinate), and 5) movement into an open area (Open Area). Descriptions of each context are listed in Table 2.2. Although grunts did seem to be produced on occasion in contexts that fell outside these 5 main categories, they did so at levels that precluded meaningful analysis (These calls are categorized as "Other" in Table 2.2). These "Other" contexts included the initiation of group movement after a period of rest (group movement not into an open area: Males; n = 1, Females; n = 8), vigilance at unknown (Males; n = 7, Females; n = 5), and social play (Males; n = 1, Females; n = 0). Additionally, there were many instances that I could not detect an obvious social or ecological context surrounding the production of a grunt. To avoid the assumption that the grunt was simply "spontaneous", I assumed I had simply missed the contexts that elicited these

calls, and labeled them as “unknown” (unable to determine context; Males; n = 40, Females; n = 26). These Unknown contexts were also lumped in with the “other” context (Table 2.2).

Table 2.2: Definitions of the different contexts in which grunts were produced.

Aggression
Grunts and grunt-like calls produced during overtly aggressive interactions (e.g., facial threats, lunges) were classified as “aggressive grunts”. Although these grunts were often produced by subordinate animals in the presence of a dominant individual, they were distinguished from the mere “presence of a dominant” context by the intensity of the interaction, involving a targeted action by one animal toward another.
Inter-Troop Encounter (ITE)
Grunts produced while the caller was looking at individuals from another nearby social group, but did not include calls given while the caller was actively engaged in aggression with the other group. These grunts could at times be emitted among other calls fitting the description of the long Aarr, long Aarr-raugh, and long rraugh calls as described by Struhsaker (1967a), which were characterized by being much longer than what would typically be described as a grunt, and consisting of more pulses. As a cutoff, I only used calls which consisted of 5 or fewer elements.
Presence of dominant individual (Dominant)
I considered calls to be elicited by the presence (or detection) of a dominant individual if the caller was clearly looking at a dominant individual (relative to themselves) when calling. I also noticed in the field that vervets grunted when in proximity to a relatively dominant individual, despite the caller not looking at the dominant animal while calling. As such, I expanded the definition of “call in the presence of a dominant” to include calls that did not meet the criteria for another context, but where the caller was within 5 meters of a dominant individual.
Presence of a subordinate individual (Subordinate)
I defined calls as being elicited by the presence of a subordinate individual when grunts were produced by a caller who approached and/or displaced a subordinate monkey. Grunts produced in this context were relatively rare, which is consistent with observations made by Cheney and Seyfarth (1982a), who also reported that this contextual use of grunts was less common than other contexts in their analyses.
Group travel: Moving into an open area (Open Area)
I defined an open area as one where the average distance between the base of shrubs and trees exceeded 5m, and there was little to no canopy coverage overall. Grunts produced while the majority of group members initiated travel or were moving into an open area were placed into this category if the caller(s) was not also within 5 m of a dominant individual (or obviously looking at one).
Other
Grunt vocalizations that did not fit any of the descriptions above. For instance, I recorded one instance of an adult male producing grunt-like vocalizations while playing with a juvenile. Grunts were also sometimes produced when a group of resting monkeys began to move again (but not into an open area), or while vigilant towards an unknown target. Unfortunately, calls in these categories (i.e., play, vigilance, begin movement) were so uncommon that they precluded meaningful analysis (see text above), and are not addressed in the contextual analysis. I also recorded grunts that appeared to be produced spontaneously, outside of any discernable context. I did not enter “spontaneously” as a context for these “unknown” contexts, as I did not feel I could be sure that these grunts truly were spontaneous; it was possible that as an observer I had missed some relevant social or ecological event surrounding the call’s production.

Selecting calls for analysis

Like many vocal classes in primates, vervet grunts have been variously described by researchers, often subjectively and onomatopoeically. In their original paper, Cheney and Seyfarth (1982a) described grunts as being low pitched, pulsatile vocalizations, corresponding to the ‘progression grunt’ and the ‘woof’ described by Struhsaker (1967a). While I tried to follow these descriptions as a guide in selecting grunts from my recordings, I also found that vervets produced additional grunt-like vocalizations when encountering another social group, that seemed to correspond to the short “Aarr”, short “Aarr-rraugh”, and short “Rraugh” calls described by Struhsaker (1967a). Similar to the grunts described by Cheney and Seyfarth (1982a), these calls were relatively short, low pitched, and often consisted of multiple pulses. Subjectively, these calls seemed similar to grunts emitted in other contexts, and I only identified them as being analogous to Struhsaker’s short Aarr, short Aarr-rraugh, and short rraugh calls because they were emitted alongside calls fitting the descriptions of the long Aarr, long Aarr-rraugh, and long rraugh calls. The “long” and “short” versions of these calls differed primarily in their length, and in the number of elements (see above) that made up each call (Struhsaker 1967a).

I selected calls after listening to the recordings, and visually examining their spectrograms. I only selected calls for analysis when at least one call within a bout was relatively free of background noise which could affect spectral analysis of the signal (i.e., other animals calling, noises from branches moving, human vocalizations), and where the start and end of the signal was clearly distinguishable from other sounds. A breakdown of the number of grunts analyzed for each age sex class is presented in table 2.3.

Table 2.3: Total number of grunts analyzed for each age sex class

	Females	Males	Total Per age class
Yearlings	73	51	124
Juveniles	172	85	257
Subadults	157	95	252
Adults	330	370	700
Total per sex:	732	601	Total: 1333

Individual grunts, and grunt-like vocalizations were often produced in bouts of multiple calls. Most calls could also be broken down into smaller elements separated by breaks (gaps, characterized by energy levels approximating background noise) in the call. I defined an element as a continuous tracing on an oscillogram whose energy was above that of the ambient background noise (Fig. 2.1). I defined a single call as a cluster of elements whose inter element time(s) were less than the amount of time between other similar calls. I defined a bout as a cluster of one or more similar calls emitted under one clear context, and separated from other bouts by a time interval longer than any time interval between calls within the bout. Because capturing entire bouts of calling was often impractical, I conducted all analyses at the level of the call and the call element only, and I do not examine potential variation in call bouts.

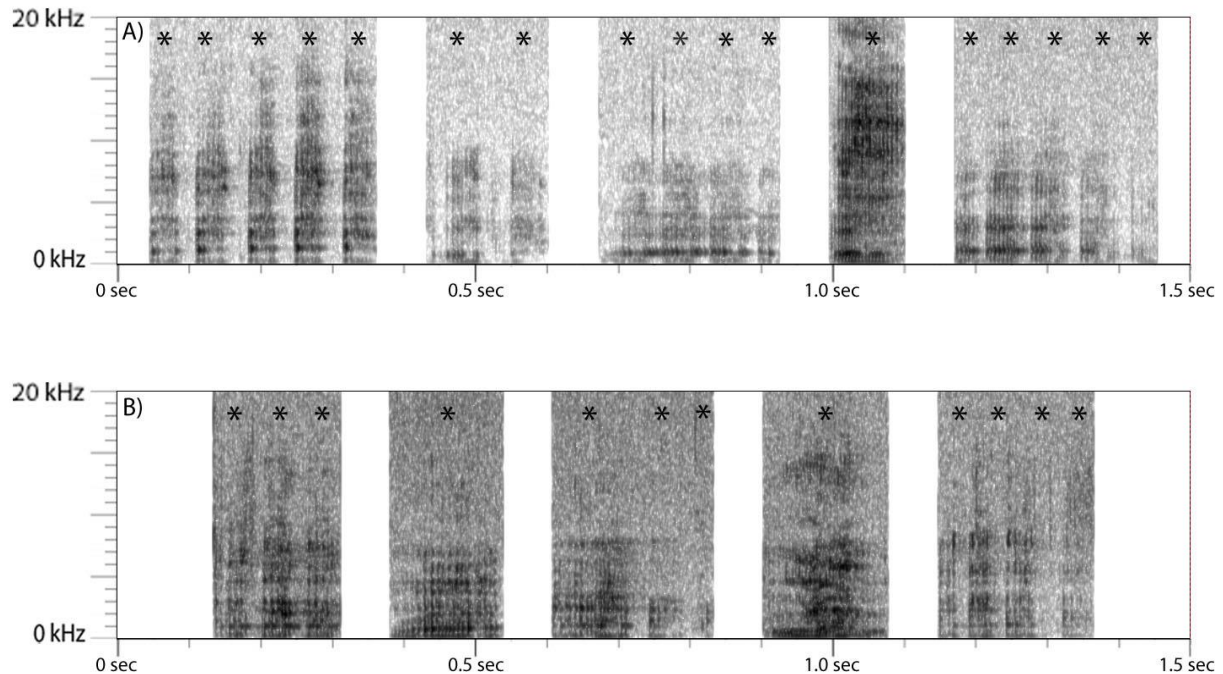


Figure 2.1: Spectrograms representing grunts A) produced by females, B) produced by males. Individual elements for each call are marked with an Asterisk (*). The entire segment for each set of calls is 1.5 seconds long.

Acoustic analysis

Temporal analysis

I made temporal measurements using a combination of the waveform (oscillogram) and spectrogram views within Raven Pro - Interactive sound analysis software, version 1.5 (Cornell Laboratory of Ornithology, Ithaca). Temporal measurements at the element and call level included the duration of the call, the number of elements within the call, and the mean duration of the call elements that made up a call.

Spectral Analysis

I selected one call from each bout for spectral analysis. I selected the first call within the bout that was clear of background noise, and that had a high signal-to-noise ratio. Each of these calls were high pass filtered at 100Hz to remove noise below the lowest frequency of the call (Raven Pro, v1.5). Each element of the selected call was then labeled, and saved as a separate file. I used the batch processing function of Avisoft SASLab Pro to pad each element with 0.2 second silent margins, and to convert the sampling frequency of each element. For my analyses concerning caller age, sex, and identity, I resampled calls to 16kHz. For the contextual analysis, I resampled the calls to 5.5kHz. I chose this sampling frequency based on the results of Cheney and Seyfarth (1982) and Seyfarth and Cheney (1984), who found that grunts produced in different contexts could be distinguished based on variations in their spectral features between 0Hz and 1200Hz. I used SASLab to generate a spectrogram of each call element via a Fast Fourier transform (Hamming window function, 1024 points, 93.75% overlap for analysis of caller age, sex, and ID. 96.87% overlap for analysis of context). The spectrograms were then exported and analyzed using custom sound analysis software developed by K. Hammerschmidt (LMA - Lautmusteranalyse v. 2018_0.4). In many instances, grunts lacked a harmonic structure, and I could therefore not measure fundamental frequency. Instead, I measured variables relating to the distribution of energy within each element of the call (Table 2.4).

First, I used the distribution of frequency amplitudes (DFA) as a metric for the distribution of energy within each time segment. DFA 1, 2 and 3 represent the frequencies which separate the first and second, second and third, and third and fourth quartiles of the overall energy distribution of a time segment respectively. Next, I measured the mean first dominant frequency band of the call. This was the lowest frequency within a time segment that exceeded a predefined threshold

across a series of consecutive frequency bins. In tonal calls, the first dominant frequency band is often equivalent to the fundamental frequency, though because of the non-harmonically structured nature of many of the calls in my sample, it is also possible that the first dominant frequency band represents a resonant frequency of the vocal tract. I measured the peak frequency, which was defined as the frequency with the highest amplitude in a given time segment of the call. For this analysis, I took the average of DFA 1, 2, 3, dominant frequency (DF1), and peak frequency (pfmean) across the duration of the entire element. Next, I measured the Wiener entropy across the entire duration of the element. Wiener entropy is a measure of the uniformity of the energy distribution across the spectrum. Here, a signal that is characterized by random energy across all frequencies (i.e., noise) has a Wiener entropy (mean_noise) of 1, while a pure tone would have a value of 0.

I also measured a number of acoustic variables that described the course of these frequency-related variables throughout each element of the call. First, I measured the location of both the maximum and minimum dominant frequency and peak frequency. Next, I measured the global modulation of the peak frequency and first dominant frequency, which represents the slope of the line which best represents the trend of pfmean and DF1 throughout the course of the call. Here, elements that were characterized by rises in their peak frequency or first dominant frequencies would have positive values for these variables, while calls characterized by decreases would have negative values. As a measure of the variability around the trend of the peak frequency, I measured the maximum deviation between the linear trend of the peak frequency, and the mean deviation between the linear trend of the peak frequency.

Finally, because previous research suggests that grunts emitted in different contexts vary primarily in the change in frequency content throughout the entire call (Cheney and Seyfarth

1982a, Seyfarth and Cheney 1984), I also calculated the difference in the peak frequency, DFA2, mean noise, DF1, and overall frequency range between the first and final call element. In instances where the first (and/or last) element of the call was/were too low in amplitude for analysis, I calculated the difference in these parameters between the earliest and latest elements in the call that were of sufficient amplitude. In instances where calls were only made up of a single element, I calculated the differences in these variables between the first and last half of the call.

Table 2.4: Variables measured from grunt vocalizations

Temporal measurements – Based on entire call	
Call_Duration	Duration of call [ms]
Number_of_elements	Number of call elements within a call [ms]
Element_length	Mean element length within a call [ms]
Frequency based measurements – Based on single element	
Pfmean	Mean peak frequency [Hz]
Pfmaloc	Location of the maximum peak frequency [(1/duration)*location]
Pfmiloc	location of the minimum PF [(1/duration)*location]
Pfjump	maximum difference between successive PF's [Hz]
Pftrfak	factor of linear trend of PF (global modulation)
Pftrmean	mean deviation between PF and linear trend [Hz]
DFA1	Distribution of the frequency amplitudes 1: Mean first quartile of amplitude in spectrum
DFA2	Distribution of the frequency amplitudes 2: Mean second quartile of amplitude in spectrum
DFA3	Distribution of the frequency amplitudes 3: Mean third quartile of amplitude in spectrum
mean_noise	Mean value of noise within call. 0=pure tone, 1=random noise
DF1	Mean frequency of 1 st dominant frequency [Hz]
maxloc	Location of the maximum frequency 1 st DF [(1/duration)*location]
minloc	Location of the minimum frequency 1 st DF [(1/duration)*location]
df1trfac	factor of linear trend of 1st DF (global modulation)
Frequency based – Entire call	
Change_pfmean	Difference in pfmean between the last and first segment of the call.
Change_DFA2	Difference in DFA2 between the last and first segment of the call.
Change_mean_noise	Difference in mean_noise between the last and first segment of the call.
Change_DF1	Difference in DF1 between the last and first segment of the call.
Change_range_mean	Difference in mean range between the last and first segment of the call.

Statistical Analysis

Random Forest models

I used a series of random forest models (Brieman 2001) to determine whether grunts differed between different sexes, age classes, individuals, or contexts. Random forest models are an expansion of standard Classification and Regression Tree (CART) models. CART models use a recursive partitioning approach, whereby data are partitioned repeatedly into multiple sub-sets (nodes), with the goal of attaining final sub-sets (terminal nodes) that are as pure as possible with regards to a particular outcome variable (e.g., context of calling: Kassambara 2017a). When the outcome variable is categorical, CART analysis produces a series of rules, which together form a “classification tree” that can be used to predict the outcome variable based on the variables used to construct the tree. Random forest models build on this method by generating multiple decision trees (i.e., a forest). Each tree is generated using a random selection of the independent variables at each splitting step (node) in each tree, along with $\sim 2/3$ of the actual cases. In each of my models, 500 trees were produced. The final classification of each call is based on how it was classified most commonly by all the trees within the forest.

To estimate the classification error, the accuracy of each tree was tested using the $\sim 1/3$ of the data that was not used to construct the tree at each iteration (called the out of bag “OOB” data). These OOB predictions are aggregated, and used to calculate an overall OOB error rate for the entire model (Brieman 2001, Liaw and Wiener 2002). To determine whether the random forest models generalize well to data that were not used to train them, I randomly selected approximately 70% of the data to create the model, and used the remaining 30% of the data for validation.

Random forest models can also be used to evaluate the relative importance of different independent variables in predicting the outcome variable. The “mean decrease in Gini” index is a

measure of how much the purity of the nodes of the decision trees generated by the model would decrease if a given variable was randomly permuted with regards to the outcome variable. Variables with higher values for their mean decrease in the Gini index are considered to be more important variables for the given classification problem.

Previous studies have relied on discriminant functions analysis for classifying calls based on acoustic structure (Fischer et al. 2002, Crockford et al. 2004, Meise et al. 2011.). I chose to use random forests instead of a discriminant functions analysis for several reasons. First, random forests are robust against non-normally distributed data and outliers, therefore allowing me to use my raw data, with no transformations, thus greatly increasing the interpretability of my results. Secondly, random forests allow predictor variables to be correlated. This allowed me to use all my acoustic variables in my analyses, letting the model select the most important variables.

To determine whether the random forest models were able to classify calls to their correct context at levels that exceeded chance expectation, I ran a series of permutation analyses. This involved randomly permuting the grouping variable for each call (i.e., caller sex, caller age, caller identity, context) in the sample, then comparing this permuted value to the actual value in the un-permuted data. This procedure was repeated 1000 times, generating a distribution representing prediction accuracy as expected by chance for each class. This was then compared to the actual prediction accuracy from each model.

Permuted discriminant functions analysis

As I analyzed multiple grunt vocalizations produced by the same individuals (repeated measures), I re-analyzed a subset of the data using a series of permuted discriminant function

analyses (pDFA), which allowed me to control for the influence of non-independent data within a data set (Mundry and Sommer 2007).

For the analysis of age and sex, I used a nested design, in which individuals (control factor) were randomized repeatedly (permuted) between different ages or sexes (test factors). For the analysis of context, I used a crossed design, in which the context of call production (test factor) was permuted within individuals (control factors). The randomization process should effectively mask any discriminability between the different test factors if they exist. The discriminability of the different test factors in the permuted data sets was compared to the discriminability of the test factors in the original data set, with the null hypothesis that the discriminant functions analysis is no better at discriminating between the true (non-randomized) test factors than it is at discriminating between two permuted (random) test factors. I ran all pDFAs described below using 1000 permutations.

2.4 Results

Classification of caller sex

I ran a random forest model to determine whether grunts vary in structure based on the sex of the individual producing the call. The model was able to correctly classify calls by sex in the training set 69.80% of the time (OOB estimate of error rate: 30.20%). This classification accuracy generalized well to the validation set, where calls were classified accurately 69.52% of the time (95% CIs : 0.6282, 0.7567), which was significantly better than chance expectation (permutation test; $p = 0.001$, expected correct classification by chance: 50.13%). Calls produced by both males and females were classified correctly at levels that exceeded chance expectation (Table 2.5). To

control for caller identity, I ran a permutated discriminant functions analysis (pDFA) on a subset of 8 adult females and 10 adult males from whom I had recorded the most calls (min = 14, max = 70). The pDFA revealed that in adults, while controlling for caller identity, the sex of the caller could be predicted at levels greater than chance (cross classified; $p = 0.002$; correctly cross classified 69.68% of calls, expected cross classified by chance = 56.13%).

Table 2.5: Confusion matrix - The number of correct classification and misattributions for each caller category (adult Males vs adult Females). Correct classifications are highlighted in green. Misclassification rate represents the percent of calls ascribed to each caller category that were misclassifications. For each caller category, permutation tests were used to compare the percent of calls correctly classified to the percent expected by chance.

		Actual Sex		Misclassification rate
		Male	Female	
Predicted Sex	Male	76	29	27.62%
	Female	35	70	33.33%
Classification accuracy		68.47%	70.71%	
Expected classification accuracy by chance		52.94%	47.23%	
P-value		<0.001	<0.001	

A total of six variables had Gini scores that were above the mean Gini score for all variables in the model (mean Gini index = 11.08, Fig. 2.2). These variables were the mean of the second quartile of the distribution of frequency amplitudes (DFA2), the mean element length, the mean peak frequency (pfmean), the duration of the call, the frequency of the first (lowest) dominant frequency band (DF1), and the factor of the linear trend of the first dominant frequency band (df1trfak). A series of Mann–Whitney U tests revealed that grunts produced by males had

significantly lower values for DFA2 ($U = 84630.500$, $p < 0.001$), longer elements ($U = 39652.500$, $p < 0.001$), longer call durations ($U = 71673.500$, $p < 0.001$), and lower peak frequencies ($U = 74851.500$, $p < 0.001$) when compared to grunts produced by adult females (Fig. 2.3). I found no significant sex differences between the average DF1 ($U = 59931.500$, $p = 0.675$), or pptrfak ($U = 64697.500$, $p = 0.172$).

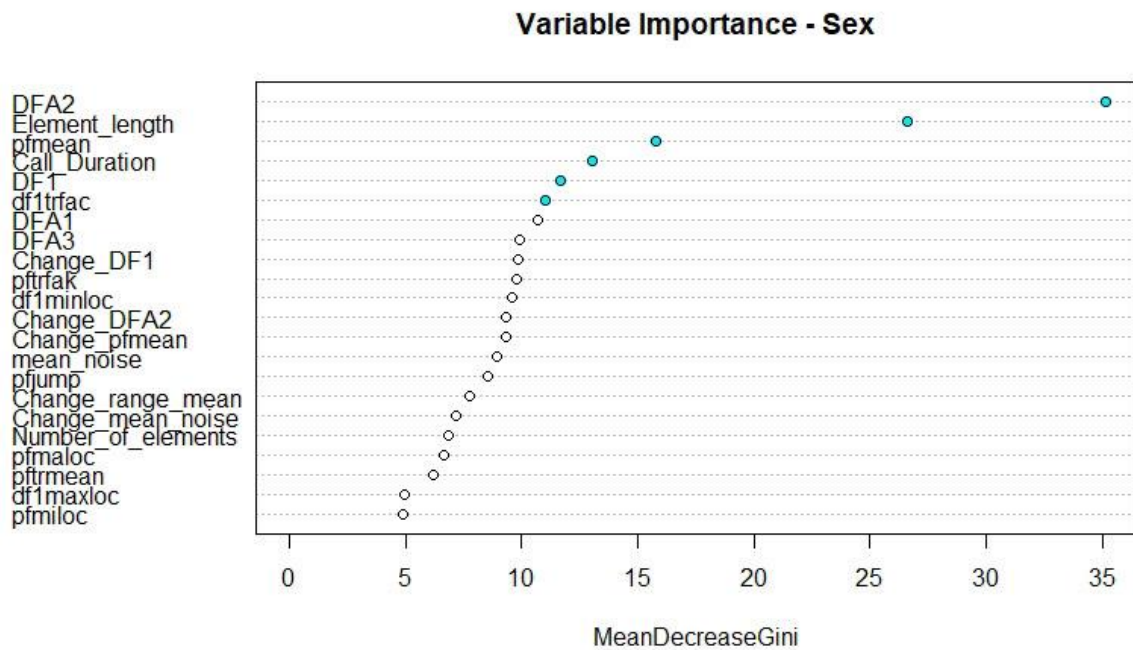


Figure 2.2: Variable importance plot for predictor variables from random forests classifications used for predicting caller sex. The “mean decrease in Gini” index is the average decrease in node impurity that results from splits using a given acoustic variable. Higher values of mean decrease in Gini indicate variables that are more important to the classification. Cyan dots represent Gini indices higher than the mean Gini index for all variables in the analysis. Descriptions of each acoustic variable are provided in Table 2.4.

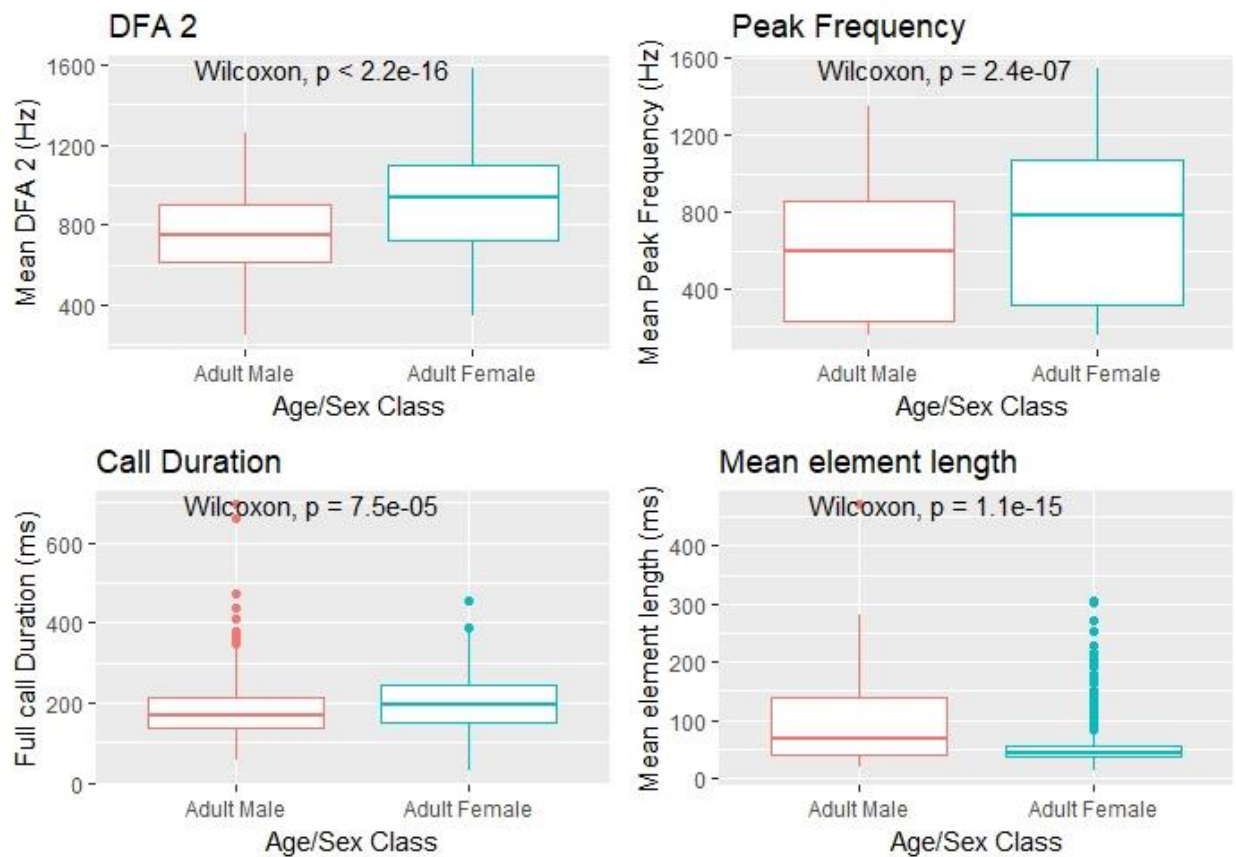


Figure 2.3: Boxplots representing the distribution of four variables which varied significantly between adult males and adult females; DFA2, Peak frequency, Call duration, and Mean element length. In each plot, the center line represents the median. The lower and upper hinges (bottom and top of each box) represent the first and third quartiles of the data respectively. Upper whiskers extend to the largest value that is no further than 1.5 times the inter quartile range (distance between the first and third quartiles) from the upper hinge. The lower whiskers extend to the smallest value that is no more than 1.5 times the inter quartile range from the lower hinge. Points outside the whiskers represent outliers.

Classification by age class

I ran a second random forest model to determine whether grunts vary in structure based on the age class of the caller. I analyzed a total of 1333 grunts produced by individuals across the four age classes (see Table 2.3). The random forest model had an accuracy of 70.24%, and an OOB error rate of 29.76%. Data from the validation set were classified correctly 71.18% of the time

(95% CI: 0.665, 0.756) which was significantly better than chance (permutation analysis, $p < 0.001$, expected classification by chance = 35.81%). Each age class was classified correctly at rates that exceeded chance expectation (Table 2.6). Importantly, where errors occurred, they did not seem random; callers from each age class were often misclassified as being produced by a caller from the age class either directly above, or below the actual age class of the caller (see Table 2.6). A subsequent pDFA which controlled for caller identity confirmed that grunts could be classified to their correct age class at levels that exceeded chance (36.92% vs 23.24% $p < 0.001$)

Table 2.6: Confusion matrix - The number of correct classifications and misattributions for calls produced by different age classes. Correct classifications are highlighted in green. Misclassification rate represents the percent of calls ascribed to each caller category that were misclassifications. For each caller category, permutation tests were used to compare the percent of calls correctly classified to the percent expected by chance.

		Actual Age Class				Misclassification rate
		Yearling	Juvenile	Subadult	Adult	
Predicted age class	Yearling	14	8	1	5	50.00%
	Juvenile	12	45	14	6	41.56%
	Subadult	4	12	38	12	42.42%
	Adult	7	12	22	187	17.98%
Classification accuracy		37.84%	58.44%	50.68%	89.05%	
Expected by chance		9.28%	19.27%	18.75%	52.63%	
P-value		<0.001	<0.001	<0.001	<0.001	

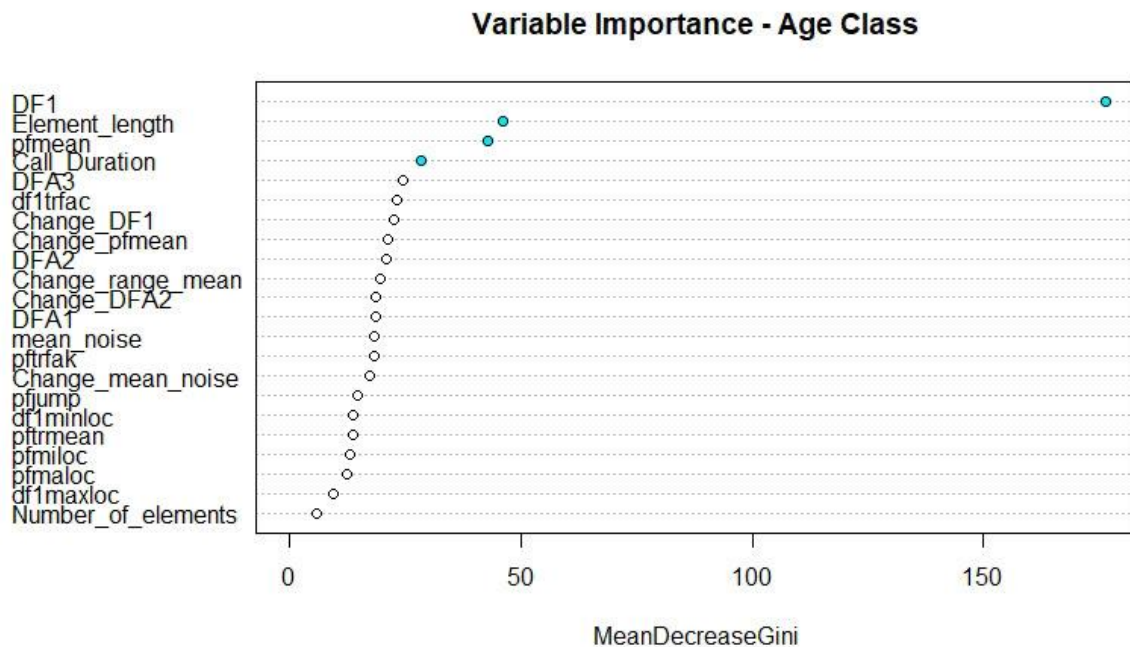


Figure 2.4: Variable importance plot for predictor variables from random forests classifications used for predicting caller age category. The “mean decrease in Gini” index is the average decrease in node impurity that results from splits using a given acoustic variable. Higher values of mean decrease in Gini indicate variables that are more important to the classification. Cyan dots represent Gini indices higher than the mean Gini index for all variables in the analysis. Descriptions of each acoustic variable are provided in Table 2.4.

Of the 22 variables entered into the analysis, only 4 variables were rated as having Gini indices higher than the average Gini index for all variables (Mean Gini index = 27.31); DF1, mean element length, Pfmean, and the full duration of the call (Fig. 2.4). Subsequent Kruskal-Wallis analyses revealed significant differences in all four variables between callers of different age classes: DF1 ($X^2(3) = 598.39$, $p < 2.2e-16$), pfmean ($X^2(3) = 12.65$, $p = 0.005$), Call Duration ($X^2(3) = 93.23$, $p < 2.2e-16$), Element length ($X^2(3) = 177.91$, $p < 2.2e-16$). As a general trend, DF1 decreased as individuals aged, while the duration of grunts, and the mean length of the

elements that made up each call increased with age (Fig. 2.5). Although pfmean seemed to decrease between yearlings and juveniles, it then appeared to increase as individuals grew older.

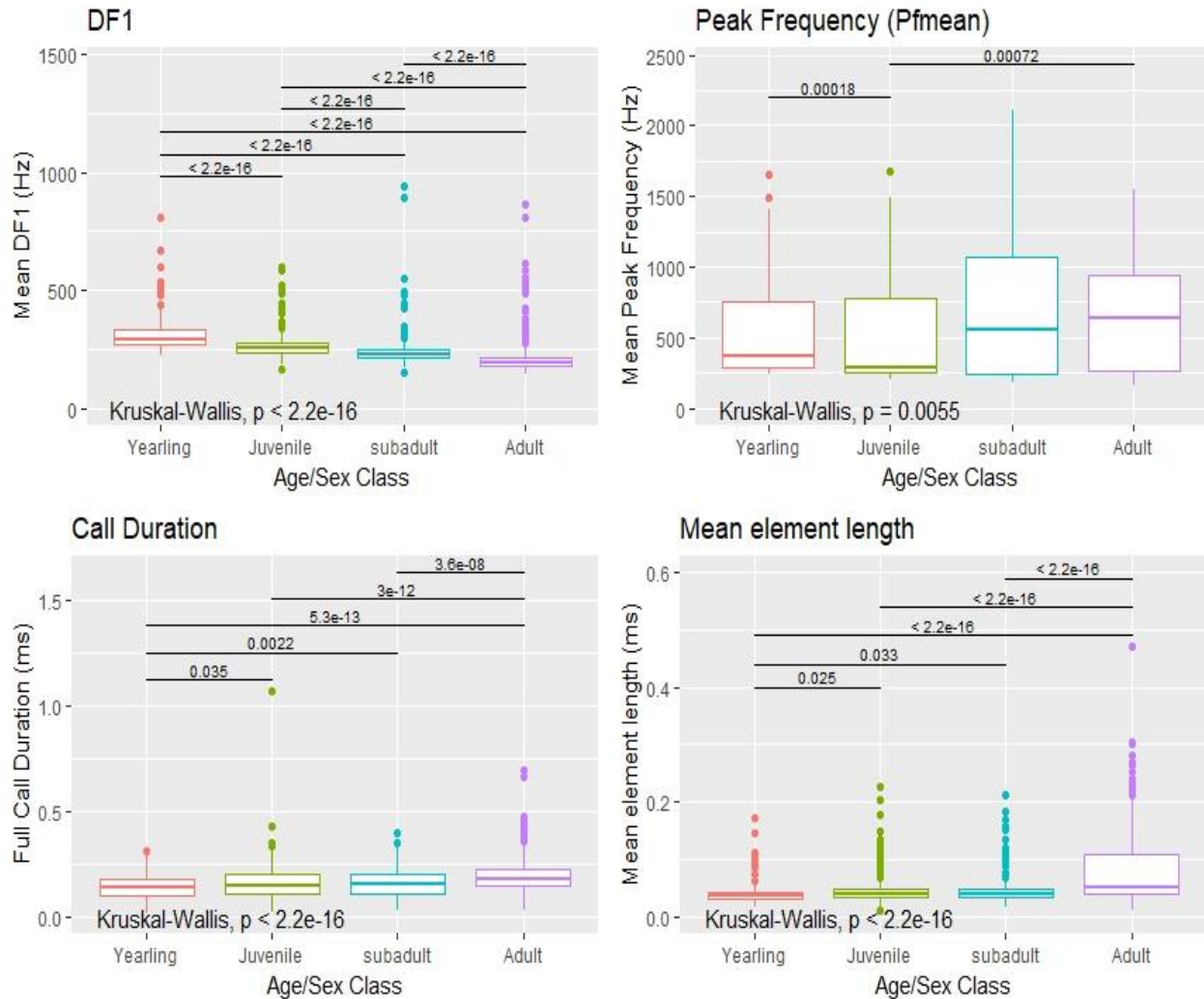


Figure 2.5: Boxplots representing the distribution of four variables which varied significantly among age classes; DF1, Peak frequency, Call duration, and Mean element length. In each plot, the center line represents the median. The lower and upper hinges (bottom and top of each box) represent the first and third quartiles of the data respectively. Upper whiskers extend to the largest value that is no further than 1.5 times the inter quartile range (distance between the first and third quartiles) from the upper hinge. The lower whiskers extend to the smallest value that is no more than 1.5 times the inter quartile range from the lower hinge. Points outside the whiskers represent outliers.

Classification of individuals using their grunt vocalizations

To test whether grunts varied in structure predictably among different callers, I ran a random forest model on grunts recorded in the “dominant” context, as this was the context for which I had recorded the most calls from both males and females. I ran the analysis on the six individuals (three adult males and three adult females) from which I had recorded the most grunts in this context (min = 12, max = 27). The random forest model had an out of bag (OOB) error rate of 60.00%, and an overall accuracy of 40.00%. Classification accuracy for the validation set was 54.35% (95% CIs: 0.3901, 0.6910), which was significantly better than chance expectation ($p < 0.001$, Expected accuracy by chance = 18.72%). Still, this classification accuracy was highly variable across different individuals (max = 100%, min = 0%), suggesting that while some individuals may produce grunts that are acoustically distinct from other individuals, this is not the case for all individuals (Fig. 2.7). Additionally, misclassification rates were relatively high for most individuals, including some with high prediction accuracies (e.g., AF1, see Table 2.7), meaning that a large number of calls predicted to have been produced by those callers were actually misclassifications. Overall, the model seemed to suggest that grunts vary in structure reliably among some individuals, though, the inability of the model to correctly classify calls produced by all individuals suggests that grunts, at least those emitted in the “dominant” context, may not be the most accurate source of information regarding caller identity.

Table 2.7: Confusion matrix - The number of correct classifications and misattributions for calls produced by each individual in the analysis. Correct classifications are highlighted in green. Misclassification rate represents the percent of calls ascribed to each caller that were misclassifications. For each individual caller, permutation tests were used to compare the percent of calls correctly classified to the percent expected by chance.

		Actual Caller						Misclassification rate
		AM1	AM2	AM3	AF1	AF2	AF3	
Predicted caller	AM1	1	0	1	0	0	0	50.00%
	AM2	1	6	2	0	0	1	40.00%
	AM3	1	1	7	0	1	1	36.36%
	AF1	5	0	0	4	0	2	63.63%
	AF2	0	2	0	0	7	0	22.22%
	AF3	1	2	0	0	0	0	100%
Percent correctly classified		11.11%	54.54%	70.00%	100%	87.50%	0%	
Expected by chance		19.66%	23.50%	21.70%	9.08%	17.25%	8.58%	
Sig.		0.548	0.003	<0.001	<0.001	<0.001	0.298	

Effects of context on grunt variation

I analyzed a total of 700 grunts produced by adult males and females across five contexts (Table 2.8). Despite efforts to sample males and females equally across contexts, a chi-square analysis revealed significant sex-differences in the contexts in which grunts were typically

produced in my sample ($X^2 = 175.4$, $df = 4$, $p < 0.001$). Females rarely produced grunts in response to subordinates or when moving into open areas, and males rarely produced grunts in the ITE (Inter Troop Encounter) context (Table 2.8).

Table 2.8: Contextual breakdown of Grunts by adult males and adult females. Expected values are indicated in brackets.

	Dominant	ITE	Subordinate	Aggression	Open Area	Unknown/ Other	Total per Sex class
Adult Male	144 (131.09)	5 (48.63)	78 (51.27)	38 (53.39)	56 (39.11)	49 (46.51)	370
Adult Female	104 (116.91)	87 (43.37)	19 (45.73)	63 (47.61)	18 (34.89)	39 (41.49)	330
Total per context	248	92	97	101	74	88	Total: 700

As my sample indicates that males and females varied in the degree to which they produced grunts in different contexts, I ran separate contextual analyses for males and females, omitting contexts for which sample sizes were too small to permit meaningful analyses for each sex. Because my sample size for calls produced in “Subordinate”, and “Open area” contexts were relatively small for adult females, my contextual analysis for females only made use of calls produced in “Dominant”, “ITE”, and “Aggressive” contexts. Conversely, because my sample size for calls produced by males in the “ITE” context was relatively small, I ran statistical analyses related to context for males with only the grunts emitted in “dominant”, “subordinate”, “aggressive”, and “open area” contexts.

Females

Because the sample was unbalanced regarding the three contexts, I used the “caret” package (Kuhn et al. 2018) to oversample the contexts that were underrepresented in the training data set (Japkowicz 2000). The model’s OOB error rate was relatively low for the training data set (OOB error rate = 33.86%, accuracy = 66.14%). This did not generalize over to the validation data set, however, which had a prediction accuracy of 39.00% (95% CI: 0.294, 0.4927), which was not significantly better than chance expectation (predicted accuracy expected by chance 34.66%, $p = 0.110$). When looking at each context separately, neither calls produced in the dominant, or ITE contexts were predicted accurately at rates that exceeded chance expectation (Table 2.9). Although calls produced in aggressive contexts were predicted accurately at rates that were higher than chance expectation (32.00%, permutation test, $p = 0.014$, Expected classification by chance: 25.15%), there was also a high rate of false positives for this context (53.85%), meaning that the majority of calls that were classified as aggressive grunts were actually produced in other contexts. Overall, these results suggest that grunts produced by females did not show clear evidence of context specific variation in structure.

Table 2.9: Confusion matrix: The number of correct classifications and misattributions for calls produced by females across three contexts. Correct classifications are highlighted in green. Misclassification rate represents the percent of calls ascribed to each context that were misclassifications. For each context, permutation tests were used to compare the percent of calls correctly classified to the percent expected by chance.

		Actual Context			Misclassification rate
		Dominant	ITE	Aggressive	
Predicted context	Dominant	18	15	6	53.85%
	ITE	11	13	11	62.86%
	Aggressive	12	6	8	69.23%
Percent correctly classified		43.90%	38.24%	32.00%	
Expected by chance		41.03 %	33.98%	25.15%	
P-value		0.135	0.502	0.014	

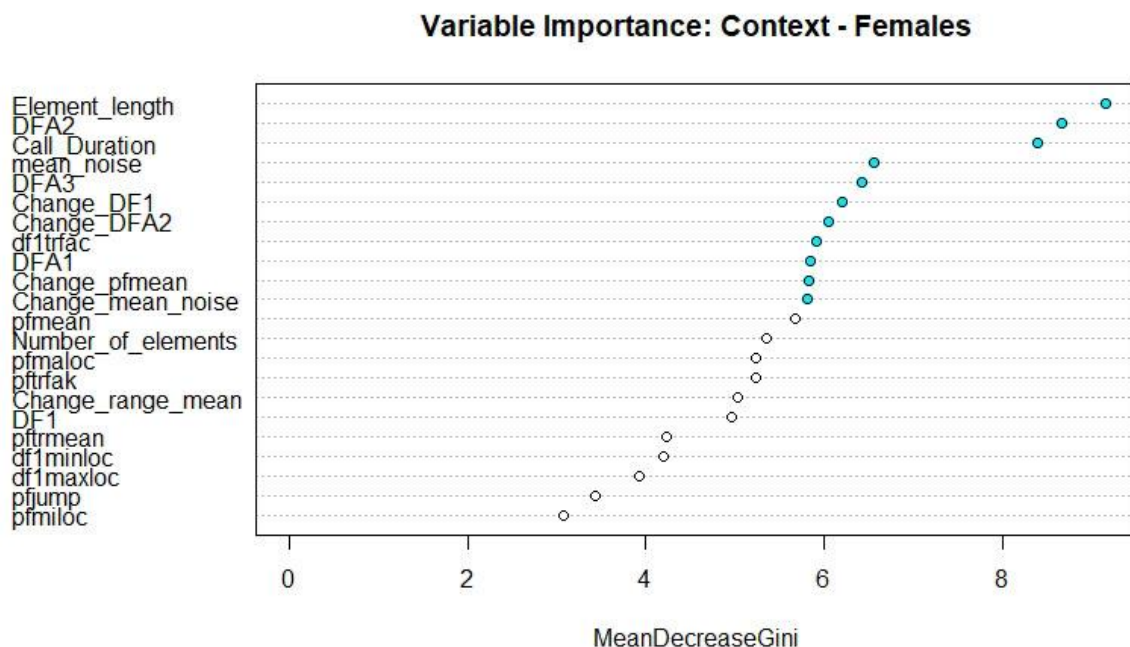


Figure 2.6: Variable importance plot for predictor variables from random forests classifications used for predicting the context in which grunts were produced (by females only). The “mean decrease in Gini” index is the average decrease in node impurity that results from splits using a given acoustic variable. Cyan dots represent Gini indices higher than the mean Gini index for all variables in the analysis. Higher values of mean decrease in Gini indicate variables that are more important to the classification. Descriptions of each acoustic variable are provided in Table 2.4.

Of the 22 acoustic variables entered into the model, 11 were classified as having Gini importance indices higher than the mean Gini index for all variables combined (mean Gini index = 5.69, Fig. 2.6). Subsequent Kruskal-Wallis analyses revealed that 5 of these variables differed significantly between contexts. These variables were Mean element length (16.60, $p < 0.001$), DFA2 (15.25, $p < 0.001$), Call duration (17.34, $p < 0.001$), DFA3 (12.93, $p = 0.002$), and DFA1 (11.148, $p = 0.004$). Pairwise comparisons revealed that mean element length was significantly longer in aggressive contexts compared to both the dominant and ITE contexts (Fig. 2.7). Call duration however was significantly shorter in aggressive contexts than it was for either ITE or dominant contexts (Figure 2.7). Both DFA2 and DFA3 were significantly higher in ITE contexts

compared to both aggressive and dominant contexts (Fig. 2.7). Calls produced in aggressive contexts also had significantly lower values for DFA1 compared to calls produced in ITE contexts (Fig. 2.7).

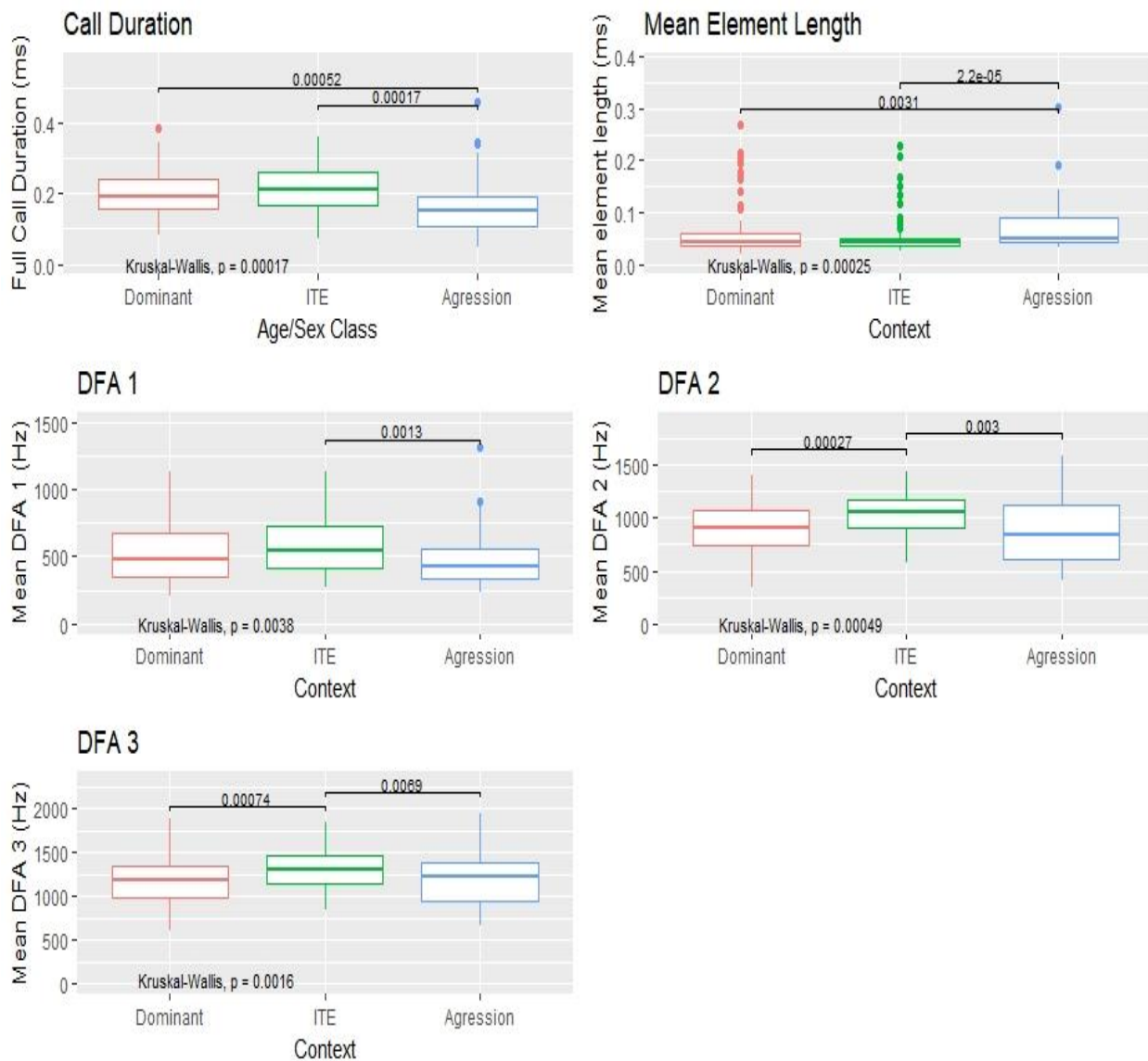


Figure 2.7: Boxplots representing the distribution of five variables which varied significantly among contexts in females; Call duration, Mean element length, DFA1, DFA2, and DFA3. In each box plot, the center line represents the median. The lower and upper hinges (bottom and top of each box) represent the first and third quartiles of the data respectively. Upper whiskers extend to the largest value that is no further than 1.5 times the inter quartile range (distance between the first and third quartiles) from the upper hinge. The lower whiskers extend from the lower hinge to the smallest value that is no more than 1.5 times the inter quartile range. Points outside the whiskers represent outliers.

I ran a crossed pDFA to determine whether the context in which calls were produced could be predicted accurately after controlling for caller identity. The crossed design of the pDFA

requires that all classes of the test factor (here, context) be paired with all classes of the control factor (here, individual ID). Because I was not able to record all monkeys calling in all contexts, I had to run the pDFA on a subset of the data consisting of only 3 adult females, from whom I had recorded grunts emitted in all three contexts used in the random forest model. The results of the pDFA suggested that the calls could not be correctly categorized based on their context of emission above chance expectation when controlling for call identity (correctly classified: 37.75%, expected by chance: 34.44%, $p = 0.25$).

Males

The random forest model had an OOB error rate of 16.38% (prediction accuracy = 83.62%). This high rate of prediction accuracy did not carry over to the validation data set however (prediction accuracy for validation set = 45.60%, 95% CIs: 0.3667, 0.5475), suggesting that the classification rules did not generalize well outside the training data.

Despite the low prediction accuracy of the model, it performed better than chance (permutation test, $p < 0.001$; Expected correct classification by chance: 31.48%). The classification accuracy for calls produced in the dominant context was 80.70%, which was better than chance (permutation test, $p < 0.001$; Expected correct classification by chance: 45.58%). Only 51.69% of calls predicted to be from the dominant context actually were, however, meaning that a large number of calls predicted by the model to have been produced in the dominant contexts were misclassifications (Table 2.10). None of the other contexts were predicted accurately at levels that exceeded chance expectation (subordinate contexts: prediction accuracy: 16.13%, permutation test, $p = 0.698$; Expected correct classification by chance: 24.85%; Aggressive contexts: 6.67%, $p = 0.258$, chance 12.10%; Open area: 22.73%, $p = 0.33$, 17.52%), suggesting that variation in the

acoustic structure of grunts produced by males was not sufficient to distinguish calls to their context of production.

Table 2.10: Confusion matrix - The number of correct classifications and misattributions for calls produced by males across four contexts. Correct classifications are highlighted in green. Misclassification rate represents the percent of calls ascribed to each context that were misclassifications. For each caller category, permutation tests were used to compare the percent of calls correctly classified to the percent expected by chance.

		Actual Context				Misclassification rate
		Dominant	Subordinate	Aggression	Open area	
Predicted context	Dominant	46	19	12	12	48.31%
	Subordinate	5	5	1	4	66.67%
	Aggression	2	4	1	1	87.50%
	Open area	4	3	1	5	61.54%
Percent correctly classified		80.70%	16.13%	6.67%	22.73%	
Expected classification accuracy by chance		45.58%	24.85%	12.10%	17.52%	
P-value		<0.001	0.855	0.575	0.153	

Variable Importance: Context - Males

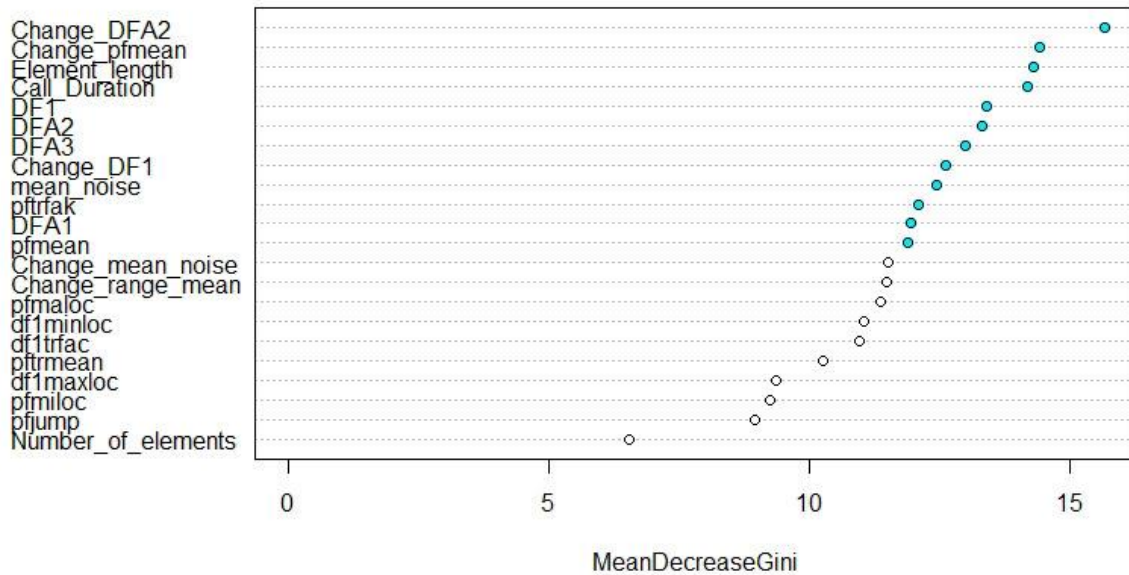


Figure 2.8: Variable importance plot for predictor variables from random forests classifications used for predicting the context in which grunts were produced (by males only). The “mean decrease in Gini” index is the average decrease in node impurity that results from splits using a given acoustic variable. Higher values of mean decrease in Gini indicate variables that are more important to the classification. Cyan dots represent Gini indices higher than the mean Gini index for all variables in the analysis. Descriptions of each acoustic variable are provided in Table 2.4.

Of the 22 variables entered into the analysis, 12 had Gini indices above the mean Gini index for the entire sample (mean Gini index = 11.82, Fig. 2.8). Subsequent Kruskal-Wallis analysis revealed significant differences in Change_DFA2 (14.01, $p = 0.003$), in the mean element length (12.93, $p = 0.005$), the full call duration (20.42, $p < 0.001$), and in the mean noise (11.80, $p = 0.008$, Fig. 2.9).

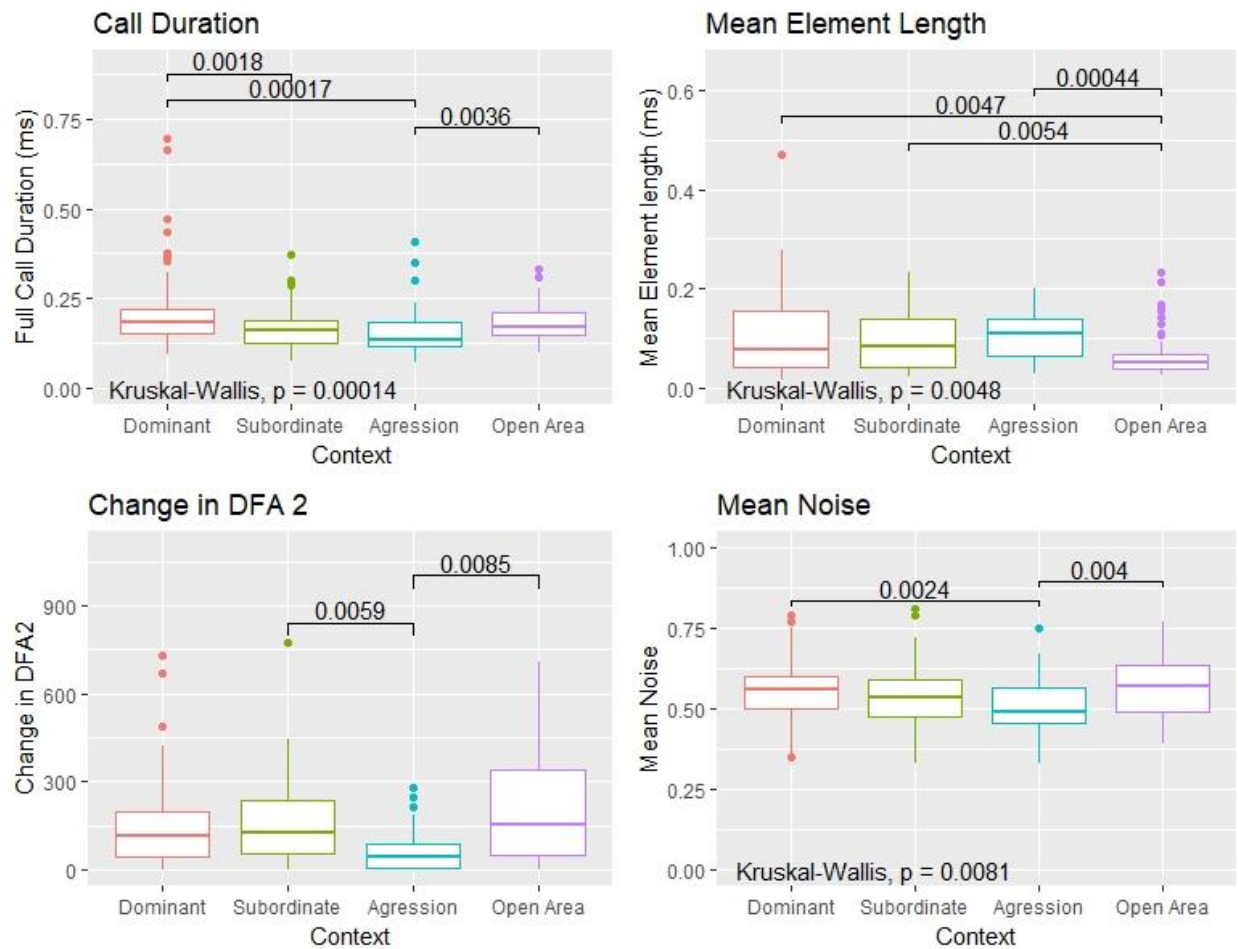


Figure 2.9: Boxplots representing the distribution of four variables which vary significantly between contexts in males; Call duration, Mean element length, Change in DFA2, and Mean noise. In each box plot, the center line represents the median. The lower and upper hinges (bottom and top of each box) represent the first and third quartiles of the data respectively. Upper whiskers extend to the largest value that is no further than 1.5 times the inter quartile range (distance between the first and third quartiles) from the upper hinge. The lower whiskers extend from the lower hinge to the smallest value that is no more than 1.5 times the inter quartile range. Points outside the whiskers represent outliers.

As was the case with the female grunts, the limited number of calls that I had recorded in some contexts, in addition to the fact that not all the subjects produced grunts in all of the contexts precluded me from using my full sample in the crossed pDFA design. As such, I was only able to

analyze grunts given by seven adult males in the “dominant”, “subordinate”, and “open area” contexts. The pDFA could not distinguish between calls produced in the three contexts at levels that were above chance (correctly cross classified: 26.63%, expected percent correctly cross classified: 25.03%, $p = 0.365$).

2.5 Discussion

Age and Sex

My results indicate that the acoustic structure of grunts varies with the age and sex of callers. Overall, these findings are in line with my predictions, which were based on what is known about the mechanisms of vocal sound production in mammals. As predicted, I found that the variables that best distinguish between grunts produced by different age- and sex-classes coincide with those that are predicted to vary based on differences in caller body size. For example, both call length and mean element length were identified as important variables for classifying grunts by both age and sex. Both these variables are likely to vary based on the size of a caller’s lungs. Larger bodied animals produce longer vocal signals to the extent to which greater lung volume provides airflow for longer periods during phonation (Fitch and Hauser 2003, Hinds and Calder 1971).

Calls produced by males and females could also be distinguished through variables relating to the spectral profile of the call, including the frequency that bisected the energy content of the signal (DFA2), and the mean value of the first dominant frequency band (DF1). Similarly, calls produced by different age classes were distinguished based on variation in their peak frequencies (pfmean), and the frequency of their lowest dominant frequency band (DF1). Larger-bodied animals are likely to possess longer/thicker vocal folds, and longer supra-laryngeal vocal tracts

than smaller individuals (Fitch and Hauser 2003). Both of these features will result in lower, or ‘deeper’ spectral profiles, as was observed here.

Despite the apparent age and sex related differences in call structure, it should be noted that the random forests models made a number of classification errors when classifying calls by both caller sex (69.52% accuracy), and age (71.18% accuracy). One interpretation of this is that there has not been strong selection on signalers to advertise their specific age or sex through grunt vocalizations. Importantly, the interaction between the physical characteristics of an individual’s vocal tract and the sound it produces can result in signals that provide honest information about signalers, even in the absence of any selective pressure to do so (Fitch and Hauser 2003). This being the case, the observed age sex differences in call structure could be considered more of a by-product of variations in body size than a feature that has resulted from selection for those traits. This is not to say, however, that this variation is not important to receivers. Sex, age, and body size in general are likely to play an important role in many forms of social interaction, including aggression, affiliation, and reproduction. The ability to accurately perceive these traits would likely prove beneficial, as, for example, an animal’s ability to defend/monopolize resources from conspecifics is likely tied to these variables (Morton 2017). As such there may be strong selection pressure for receivers to be sensitive to acoustic cues to these traits, regardless of whether there is strong selection on signalers to send this information. In this sense, age and sex related variation in grunt structure could represent *cues* as described by Bradbury and Vehrencamp (2011); assessable regularities between some aspect of a behaviour or event (in this case, the acoustic structure of grunts) and some other specific condition (age/sex), that can affect the behaviour of other organisms, but that has not evolved specifically to do so. Future studies should concentrate

on the response of receivers to variations in caller age and sex to determine whether receivers make use of these cues.

One area highlighted by my study that merits further investigation relates to sex differences in the contexts that elicit grunts in adults. My sample of grunts was unbalanced regarding the contexts in which males and females produced grunts. Based on my own observations in the field, I feel that these sex differences represent actual sex differences in the vocal behaviour of adults. For instance, generally speaking, males did not appear to produce grunts commonly during inter-group encounters. Additionally, dominant females very rarely produced grunts when approaching subordinates. Conversely, it seemed that while subordinate males often vocalized when being approached by dominant males, vocal exchanges were fairly common between both interactants (i.e., both interactants called) in these instances. Still, the data collection methods used here do not allow me to test for these differences empirically, and future work should focus on measuring sex (and age) differences in grunt use.

Individual Identity

Despite the individually-specific differences that likely characterize the vocal tracts of different individuals, my results suggest that the individual discriminability of vervet grunts is limited. While the random forest models were able to distinguish between the calls of some individuals at levels that exceeded chance expectation, they were not able to do so for all individuals. Consequently, it seems that in many instances, receivers would not be able to extract accurate information regarding a caller's specific identity from grunts, and it is possible that selection for individual differences may not be particularly strong for this vocalization. Grunts are generally relatively low in amplitude, and are often produced by animals in close proximity to one

another. As such, it is likely that individuals are either already aware of each other's identity upon hearing a signal, or that receivers can gain this information upon hearing a grunt simply by looking at the caller. Thus, there may be relatively weak selective pressure on callers to produce grunts with clear, unambiguous cues to their identity. Mechanistically, it is possible that the low amplitude of the source energy in these signals precludes them from clearly conveying the various resonant frequencies of an individual's vocal tract. Future studies should attempt to identify specific resonant frequencies (formants) in grunts, and perhaps limit their sample to higher amplitude calls that would best expose the position of these frequency peaks.

One limitation of my study was that, due to limitations in sample size, I only attempted to identify cues to caller identity in grunts emitted in one context (the "dominant" context). It is of course possible then that grunts emitted in different contexts could contain clearer cues to caller identity than were found by my analysis. For example, grunts emitted in both aggressive interactions or when another social group has been encountered (which at times foreshadow aggressive interactions between groups) could, as has been observed in the screams of other species (Cheney 1977; Gouzoules et al. 1984; Bernstein and Ehardt 1985; Gouzoules et al. 1985, 1986), serve a recruitment function, whereby caller's are able to solicit aid from close associates. This being the case, it is possible that there could be stronger selection pressure for grunts emitted in these contexts to contain cues to caller identity. Future studies with larger sample sizes should focus on this possibility.

Context

The results of my contextual analysis revealed a low classification accuracy for grunts produced in different contexts by both males (45.60% accuracy) and females (39.00% accuracy).

In contrast with the findings of Cheney and Seyfarth (1982a), this suggests that the grunt vocalizations produced by vervet monkeys do not meet the criterion of production specificity that is considered to be a prerequisite for vocal signals to act in a functionally referential manner (Macedonia and Evans 1993).

There are a number of ways in which to interpret these results. First, as already mentioned, grunts are generally produced in close range interactions between individuals. As a result, vervet monkeys are likely able to make use of visual cues when interpreting or responding to grunts in most contexts, mitigating the need for specific grunt variants that “reference” a particular context of production. Additionally, unlike the vervet’s alarm calls, grunts are not produced in contexts in which misinterpreting the ‘meaning’ of a signal is likely to have dire consequences for senders or receivers. Researchers have suggested that the evolution of distinct, functionally referential alarm calls, for example, is linked to the high potential costs of any ambiguity in the specific ‘referent’ of a given alarm type (i.e., predator type), in addition to the incompatibility of the different response types required for different predator classes (Macedonia and Evans 1993). Grunts differ substantially from alarm calls in this regard, as it is likely less crucial that grunts provide clear contextual information to receivers, as the contexts that elicit grunts are clearly less urgent. Additionally, the contexts in which grunts are generally produced may not require such specific responses, and it is unlikely that the “appropriate” response to grunts given in different contexts are in any way “incompatible” with one another; The playback experiments performed by Cheney and Seyfarth (1982a) showed that responses to grunts produced in different contexts differed in terms of the direction of the subjects’ glances relative to the position of the caller, or the duration of a given response. These results do not indicate that vervets responded to different grunt ‘variants’ with mutually exclusive responses. Instead, vervets responded by searching for

additional contextual cues to clarify the specific eliciting stimulus. As such, it is likely that the most appropriate response to a grunt may simply be to search for more contextual information surrounding its production.

The importance of contextual cues in disambiguating signal ‘meaning’ have been highlighted by a number of authors (Rendall et al. 1999, Scarantino 2010, Wheeler and Fischer 2012, Ducheminsky et al. 2014, Scarantino and Clay 2015). For instance, using playback experiments, Rendall et al. (1999) highlighted the fact that, although baboon grunts vary in structure between some contexts (i.e., moving into an open area, and approaching a female with her infant), the response of receivers to different grunt variants is highly dependent on the context in which they are perceived. For instance, while baboons seemed to respond to playbacks of different grunt types with functionally distinct responses when moving, the response of subjects did not differ between grunt variants when the group as a whole was resting. In resting contexts, baboons often showed no obvious response to either grunt variant, or alternatively, they looked towards the playback speaker as if attempting to search for additional information surround the call’s production. Similarly, recent work exploring the responses of vervet monkeys to their own species-typical terrestrial and avian alarm calls have shown that even in more ‘urgent’ contexts involving predator encounters, the modal response of receivers to alarm calls is to look towards the caller, again suggesting that the most appropriate response for receivers is to seek additional information surrounding the production of a call (Ducheminsky et al. 2014). Price and Fischer (2014) found that the response of green monkeys (*Chlorocebus sabaeus*) to their own species typical alarm calls is affected by contextual cues. Specifically, the duration of subject’s anti-predator responses to playbacks (the length of time subjects spent in the safety of trees after hearing

a playback) was affected by whether or not they had been primed with a leopard model before hearing the playback.

As the acoustic structure of vervet grunts do not seem to vary reliably based on their context of production, future research could explore whether receiver response to these signals is at all dependant on the contexts in which the signal is perceived. For instance, it was my impression in the field that vervet monkeys often encountered other social groups at some locations in their home range more than others (e.g., near territorial boundaries, especially near small contestable resources, such as pools of water). It may be the case then that grunts produced in these locations are more likely to elicit searching behaviour for extra group individuals than grunts produced in the center of the home range. Similarly, it is also the case that some contexts are associated with other call types. For example, aggressive interactions are often associated with screams and barks, while intergroup encounters are associated with intergroup ‘wrrs’ (Cheney and Seyfarth 1988; Struhsaker, 1967a). As such, it is likely that receiver response could be dependant on the integration of multiple signal types.

The differences in my results versus those observed by Cheney and Seyfarth (1982a) may be due in part to methodological differences between the two studies. First, I incorporated grunts produced during aggressive interactions, which were excluded from their original analysis. Additionally, I included grunts that were produced in bouts, whereas the original study only used grunts that were emitted as single calls. When assessing contextual variability in call structure, systematically omitting calls produced in some contexts risks over estimating how well the remaining contexts can be accurately predicted based on call structure alone (Meise et al. 2011). Thus, I attempted to explore the full variation in grunt structure that vervet monkeys produce and

are exposed to under natural conditions, including calls produced in aggressive contexts, and in bouts of multiple calls.

The methodological differences between the two studies (Cheney and Seyfarth 1982a, and the current study) do highlight some potential confounding factors in my analysis. First, although I analyzed calls produced in bouts, my acoustic analysis was carried out on individual calls within each bout. As such, it is possible that some aspect of the bout itself could vary between different contexts, which would have been missed by my analysis. It was my impression in the field for instance, that grunts produced during aggressive interactions were produced in rapid succession, while grunts produced as individuals approached one another during benign interactions were more often given as single calls, or at least at lower rates. Due to the fact that bouts were very often obscured by the calls of multiple individuals, or other vocal types (e.g., screams during bouts of aggression, or intergroup wrrs during ITE contexts), I could not reliably pull apart all grunts produced by a given individual within a bout, meaning I had to focus on single calls for my analysis. Despite this limitation, previous work on this species (Cheney and Seyfarth 1982a) was able to find contextual differences in the structure of individual grunts, suggesting contextual variation should be evident without needing to consider differences related to the bout itself. At the same time, the inclusion of grunts produced in bouts brings up the possibility that individual grunts vary in structure within bouts. For instance, grunts produced during aggressive interactions are likely to vary in structure based on whether they are produced before, or after the caller was (for instance) bitten by an attacker. As a rule, I selected the first call in any bout that was relatively clear of background noise for acoustic analysis. Because of noise generated by other monkeys, or other field workers however, the relative position of the call within a given bout varied a great deal between recordings. With this in mind, it is possible that grunts earlier in a bout may serve an

“attention grabbing”, or “priming” function, and that contextual nuances are conveyed by grunts later in the bout, which would have confounded my results. Additionally, grunts produced later in a bout may differ in structure as, for instance, a caller’s level of arousal decreases at the end of an aggressive bout, or as it calms down after a dominant individual with intentions that are initially unknown approaches and begins to show signs of affiliation towards the caller. These potential confounding factors should be evaluated, as they could have affected the results of my study. Loughry et al. (2019) recently highlighted the fact that in Gunnison’s prairie dogs (*Cynomys gunnisoni*) alarm calls within a given bout of calling varied in both temporal and spectral parameters, based on their relative position within the call bout (calls produced early in the bout vs. those produced later), highlighting the fact that variation within a given bout of calling can be a potentially confounding factor for studies focused on acoustic variation within call types. An additional limitation of my study revolves around sample size; sample sizes for some contexts in my analysis were as low as 38 calls. It is possible that this limitation affected the result of my analyses. Future studies should attempt to replicate my results with a larger sample of calls in different contexts, and from a wider range of individuals.

If the grunts produced by vervet monkeys do not contain information “about” the specific contexts in which they were produced, what then is the function of this vocal signal? One possibility is that grunts are merely a manifestation of a caller’s internal state (motivational, or emotional state). Vocal signals have long been hypothesized to be under strong motivational, or emotional influence (Darwin 1872, Goodall 1985, Seyfarth and Cheney 2003). This view was supported by self-stimulation experiments, which showed a close link between the structure of vocalizations produced by squirrel monkeys and the relative degree of adverseness/pleasantness experienced by the caller (Jürgens 1979, Jürgens et al. 2001). Similar relationships between

internal state and vocal structure have been reported across a number of taxa (birds and mammals: Morton 1977; squirrel monkey, *Saimiri sciureus*: Jürgens 1998; mammals: Briefer 2012). Provided that the acoustic structure of vervet grunts is linked to a caller's internal state, any link between external context and the acoustic structure of a given vocal signal would break down if different contexts elicit similar internal states in signalers, or if a given context does not consistently elicit the same internal state in callers. Although I did not collect data pertaining to the motivational or emotional state of callers for this analysis, it seems conceivable that the internal states of callers could vary a great deal within and between the contexts that elicit grunts. For instance, callers may experience similar levels of anxiety when they encounter another social group, when a caller is approached by a dominant individual, or in association with overt aggression. If grunt structure is strongly influenced by caller motivation or emotional state, it is conceivable that calls produced in these contexts could converge in acoustic structure.

Within each of these contexts, the internal state of a caller is also likely to vary depending on a number of factors. For instance, in the presence of a dominant animal, a caller's internal state is likely to vary based on the degree of asymmetry in the respective dominance status of two interactants. Similarly, not all aggressive interactions escalate to the same level; aggression can range from simple visual threat displays, to overt physical aggression, which is likely to affect the emotional state of both interactants. This being the case, it is likely that calls given within a particular context vary a great deal between, or even within bouts. If call structure does vary in relation to a caller's internal state, it is quite possible that the external contexts that I ascribed to each grunt in my sample may not actually align with the internal "context" which may have been the actual motivation for calling, and which might be more relevant to the actual caller and receiver.

While a caller's current internal state may not be a strong indicator of context, it is likely a strong indicator of a caller's subsequent behavior (Smith and Harper 2003). When two individuals are in close proximity, there is likely to be a high degree of uncertainty surrounding the upcoming behavior of both interactants (Silk et al. 2000, Searcy and Nowicki 2005). Grunts could play an important role in facilitating these interactions by advertising the current motivations of callers. Such motivational frameworks have been used to interpret the function of grunts in other primate species as well. For example, stump-tailed macaques (*Macaca arctoides*) produce 'staccato grunts' as individuals approach mothers with dependent infants (Bauers 1993). Analysis of social interactions associated with the production of staccato grunts suggests that females call when they are near another female's infant, and that these grunts increase a mother's tolerance of subsequent interactions between her infant and the caller (Bauers 1993). Baboons also produce grunts in association with reconciliatory behaviors post conflict, which have also been proposed to communicate a caller's motivation to interact affiliatively with another individual post conflict (Cheney and Seyfarth 1997). A number of primate species produce close-range calls in similar contexts, and it has been suggested that these vocal signals act to emphasize the benign motivations of callers, thus facilitating affiliative associations between members of the same group (e.g., stumptail macaques, *Macaca arctoides*: Bauers 1993; chacma baboons, *Papio ursinus*: Silk et al. 2000; rhesus macaques, *Papio cynocephalus ursinus*: Silk et al. 2003).

Even in the absence of contextually-specific grunt variants, it is important to remember that under natural conditions, receivers have access to a great deal of contextual information upon hearing a grunt. Vervet monkeys have visual access to the individuals in their proximity, and both aggressive and intergroup interactions are associated with a number of visual and auditory signals which could disambiguate the 'meaning' of a grunt produced in these contexts. This, in addition

to the possibility that motivational cues are available in the grunts themselves suggests that grunts could provide receivers with a great deal of specific “information”, even in the absence of specific “referential” grunt variants. In fact, a communication system that relies on the integration of contextual cues with signal structure may be more flexible, and in some ways richer than one that consists of a limited number of grunt variants that inflexibly “stand for” a select few specific contexts (Wheeler and Fischer 2012).

Chapter 3 Sex Differences in the terrestrial alarm calls of vervet monkeys

3.1 Introduction

Sex differences in vocal production and use have been documented in a large number of primate species. In some cases, sex differences manifest as divergences in the use patterns of a particular call type by males and females (e.g., Old world Monkeys: Hohmann 1991; Black and White Colobus, *Colobus guereza*: Harris et al. 2006; Muriquis, *Brachyteles hypoxanthus*: Arnedo et al. 2010; Spider monkeys, *Ateles geoffroyi*: Dubreuil et al. 2015). In these instances, some calls are produced more or less frequently by one sex, or there may be variation between the sexes in the contexts that elicit the production of a specific call type. For example, although both male and female black and white colobus monkeys produce roar vocalizations in a number of contexts, only males produce these calls during morning choruses (Harris et al. 2006). In other instances, sex differences can include the existence of sex specific calls in a species' vocal repertoire (Gautier and Gautier-Hion 1977, Waser 1982, Poole 1994, Semple 2001, Maestripieri et al. 2005; Pfefferle et al. 2008). For example, in a number of old world monkeys and apes, females produce 'copulation calls', which may increase mate guarding behaviors from preferred males, and/or encourage mate competition (e.g., baboons, *Papio cynocephalus ursinus*; O'Connell and Cowlishaw 1994, Macaques, *Macaca sylvanus*; Semple 1998; Semple and McComb 2000; mangabeys, *Cercocebus torquatus atys*; Gouzoules et al. 1998; chimpanzees, *Pan troglodytes schweinfurthii*; Townsend et al. 2008; Pradhan et al. 2006 for review).

Sex differences in vocal behavior are likely to be related to differences in the selective pressures experienced by males and females. For example, female African elephants (*Loxodonta*

africana) produce a number of vocalizations that are not used by males, and it is thought that this may reflect the higher degree of sociability exhibited by females relative to male elephants (Poole 1994). Similarly, male muriquis (*Brachyteles hypoxanthus*) produce their species-typical contact call, the “neigh” vocalization, more frequently than females, and it is argued that this is related to the high levels of social affiliation shown by males of this species (Arnedo et al. 2010).

In a number of taxa, males produce “loud calls” that are hypothesized to function in intergroup spacing, male-male competition, and mate attraction (e.g., frogs; Kelley 2004; birds: Slater and Mann 2004, Zollinger and Brumm 2015; red deer, *Cervus elaphus*; Charlton et al. 2007a,b; fin whales, *Balaenoptera physalus*; Croll et al. 2002; Baboons, *Papio cynocephalus ursinus*: Kitchen et al. 2003, reviewed in Mitani and Stuht 1998, Snowdon 2004, Delgado 2006). For example, the “loud songs” of fin whales are produced exclusively by males, and are thought to function as male breeding displays (Croll et al. 2002). Orangutan (genus *Pongo*) loud calls are only produced by flanged males and have been shown to mediate male-male spacing (Mitani 1985), and in some instances seem to attract nearby females (Fox 2002). Flanged-male orangutans possess throat sacs, which act as resonators for these loud calls, suggesting that selection on these signals is particularly strong. Similar sex specific anatomical specializations relating to vocal production include the enlarged hyoid bone of male howler monkeys (genus *Alouatta*; Ybarra 1986, 1988), the extended nasal cavities of male proboscis monkeys (*Nasalis larvatus*), and the elongated trachea of some birds that are found exclusively in males in many species (Fitch 1999).

In several guenon species (genus *Cercopithecus*) loud calls are produced exclusively by males (Gautier and Gautier 1977). In Diana monkeys (*Cercopithecus diana*) these loud calls are often produced at the detection of predators and have thus been hypothesized to function primarily as alarm calls (Zuberbühler 2002, 2003), yet they share some important similarities with the

sexually-selected loud calls of other primate species. For instance, as is the case with many secondary sexual characteristics, males only begin producing these calls at sexual maturity (Zuberbühler 2002). Because guenon loud calls are conveyed over a great distance, it has been proposed that in addition to acting as alarm calls, they could act to deter extra-group males from approaching a caller's social group (Zuberbühler 2002, 2006).

Although little work has focused specifically on this possibility, the sexual selection hypothesis does generate a number of testable predictions. For instance, from the perspective of signal structure, calls should not only be sexually dimorphic, but should also vary in structure between different individuals, thus allowing receivers to assess specific callers based on variation in the signals they produce (Snowdon 2004). Similar sex differences have been described in the alarm calls of baboons (*Papio ursinus*), in which males produce 'wahoo' vocalizations upon detecting a predator, while females and juveniles produce 'barks' (Fischer et al. 2001a). While the Wahoo seems to serve a warning function to conspecifics regarding the detection of predators (Fischer et al. 2001b), they are also associated with male competitive displays, and vary in structure based on the rank of callers, suggesting they may function in intra-sexual selection (Kitchen et al. 2003, Fischer et al. 2004). Additionally, wahoo vocalizations contain cues to caller identity (Fischer et al. 2002), and thus may provide receivers with the ability to assess the competitive abilities of specific males (Snowdon 2004, Delgado 2006).

Vervet alarm calls

Like many forest guenons, vervet monkeys (*Chlorocebus pygerythrus*) produce multiple acoustically-distinct alarm calls (Struhsaker 1967a, Seyfarth et al. 1980a,b, Price et al. 2015). Typically, the vervet alarm system is described as consisting of three call types (i.e., snake alarms,

aerial alarms, and terrestrial alarms), with each being produced in response to a different predator class. However, the original descriptions provided by Struhsaker (1967a) actually categorized these three calls into four types: snake alarms (snake Chutter), aerial alarms (Rraup), and two different call types that were both produced in response to mammalian predators – “chirps” and “threat-alarm-barks”. According to the original descriptions, adult and subadult males produced “threat alarm barks” in response to mammalian predators, whereas adult females and juveniles of both sexes produced “chirps”. Both of these calls are described as being low frequency (Hz), high amplitude calls, although, while threat alarm barks consist of multiple units comprised of a series of exhales and inhales, chirps consist only of exhales and are “abrupt”, “short”, and “sharp sounding” (Struhsaker 1967a, pp. 310-311).

The alarm calls of vervet monkeys have come to be regarded as the canonical example of referential communication in a non-human animal (Townsend and Manser 2013, Sapolsky 2017). This has generated a great deal of follow-up research focusing on variation in the acoustic structure of calls produced across different contexts and the putative informational content such variation might carry for receivers (Owren and Bernacki 1988, Price et al. 2015). However, surprisingly little work has focused on the acoustic variation among calls associated with the same predator class. Perhaps as a result, there is some ambiguity in the literature regarding the extent of age/sex related differences that characterize the vervet monkey’s terrestrial alarm call. While Seyfarth et al. (1980) acknowledged age/sex differences in the acoustic structure of terrestrial alarm calls, they lumped the two signals into one group (“leopard alarm calls”) based on the contexts associated with their production, and on the responses they evoked from receivers. They also noted that adult females could produce calls that were acoustically more similar to the threat alarm bark typically produced by males at the detection of mammalian predators, and that they seemed more inclined

to do so when the predators were in close proximity (Seyfarth et al. 1980b). Since those original studies, most descriptions of the vervet alarm system have included a three-call system (referring to the snake, eagle, and terrestrial alarm) with no mention about the sex differences that characterize the terrestrial alarm call. Some authors have described the terrestrial alarm calls in vervets as exhibiting little sexual dimorphism (e.g., Zuberbühler 2002), yet others (e.g., Price et al. 2015) have acknowledged these sex differences in structure. This discrepancy likely stems from the fact that the acoustic differences that characterize the alarm calls of males and females have yet to be quantified systematically.

Traditionally, vocal signals have been classified by observers using subjective auditory criteria in combination with visual examinations of spectrograms and power spectra (see Fischer et al. 2017 for review). This, in conjunction with observational data about the contexts in which the different signals were produced have been used to separate a species' vocalizations into functional units (call types). These traditional methods are of course vulnerable to individual observer biases and lack of consistency across measurements (Fischer et al. 2017). Current techniques still make use of spectrograms for signal classification, but benefit from the ability to use automated feature extraction, thus reducing observer biases when measuring the actual acoustic variables (Fischer et al. 2013, Fischer et al. 2017). Additionally, it is possible to use unsupervised clustering methods to group signals based on their acoustic parameters alone, without regard to the contexts in which they were produced (Fischer et al. 2013, Wadewitz et al. 2015). One advantage of this method is that it allows researchers to determine whether different vocal signals exist as discrete units, or whether they are characterized by continuous intergradation, and are therefore not defined by distinct boundaries that delineate clear, discrete call types.

In this chapter, my aim is to quantify the degree of sexual dimorphism that exists within the terrestrial alarm calls of vervet monkeys. First, I focus on sex differences in call structure. Specifically, I use a random forest model to determine how well the acoustic structures of calls produced by adult males and females can be distinguished from one another, and to determine which acoustic parameters most clearly separate alarm calls by caller sex. Next, I use a statistical clustering technique to determine objectively whether the terrestrial alarm calls of adult male and female vervet monkeys exist as two distinct call ‘types’, or whether “chirps” and “threat alarm barks” are characterized by acoustic integration between call types. Finally, as a first step in investigating whether sex differences in terrestrial alarm calls might represent a sexually selected trait, I investigate whether calls vary in structure among individuals.

3.2 Methods

Vocal data collection in the field

I made audio recordings in the field using a Sennheiser ME67 directional condenser microphone, and a Marantz PMD661 digital field recorder. Recordings were made at a sampling rate of 48kHz, and a bit rate of 1536kbps. To reduce wind and handling noise, the microphone was equipped with a blimp windshield and shock mounting system with a pistol grip (Sennheiser MZS20-1 Combo Mount/Grip/Stand). I made recordings of alarm calls in one of two conditions – 1) opportunistically when alarm calls occurred naturally at the detection of a predator, or 2) during experimental predator mount presentations.

When alarm calls occurred under natural conditions, either myself or another trained field worker would identify the eliciting stimulus. I would then locate a calling animal and record its

vocalizations while dictating its identity into a headset microphone. Whenever possible, I would attempt to record more than one individual per alarm bout.

I supplemented my naturalistic audio recordings by conducting experimental predator mount presentations. I presented predator mounts at intervals of no fewer than 14 days apart to any given group to avoid the risk of habituating the subjects to the mounts. In preparation for a predator model presentation, I would move a stuffed caracal (*Caracal caracal*) or molded plastic owl model concealed under a blanket (grey, brown, or army green in color) in the path of a target group as it traveled, out of view of any group members. Pilot work with our study groups revealed that the vervet monkey's response to both types of mount resembled those of their response to naturalistic encounters with terrestrial predators. When I was confident that no monkeys could see me, I would uncover the mount and move away in preparation to record any vocalizations produced by the monkeys. Once the mount had been spotted, I would attempt to record as many different individuals calling as I could. This number varied a great deal between trials based on the number of individuals calling, as well as the position of individuals in trees; I often found that during both the predator mount presentations and natural predator encounters, subjects, particularly younger individuals and females, would move higher into trees before/while calling, where they would stay until the mount had been removed. Individuals became harder to identify in these conditions, and as a result, these age sex classes were considerably more difficult to record than adult males, who often moved lower in the trees as they called, and approached the mount.

3.3 Data Analysis

Selecting calls for analysis

I selected calls that were clear of background noise, and had a high signal to noise ratio for acoustic analysis. In both naturally occurring as well as experimentally-induced bouts of alarm

calling, a single individual would generally produce a large number of calls. Upon listening to my recordings, and reviewing spectrograms, I felt it would be impossible to select a single alarm call from each individual that was entirely representative of all the calls they produced in a given bout of calling. For instance, calls produced by a single individual could consist of a different number of elements (defined as a continuous tracing on an oscillogram whose energy was above that of the ambient background noise; Fig. 3.1). Additionally, some calls produced by a given individual also contained audible exhale units (Fig. 3.1), while others did not (Price et al. 2015). Because of this variation within a given alarm bout, I selected and analyzed multiple calls from the same bout for each individual.

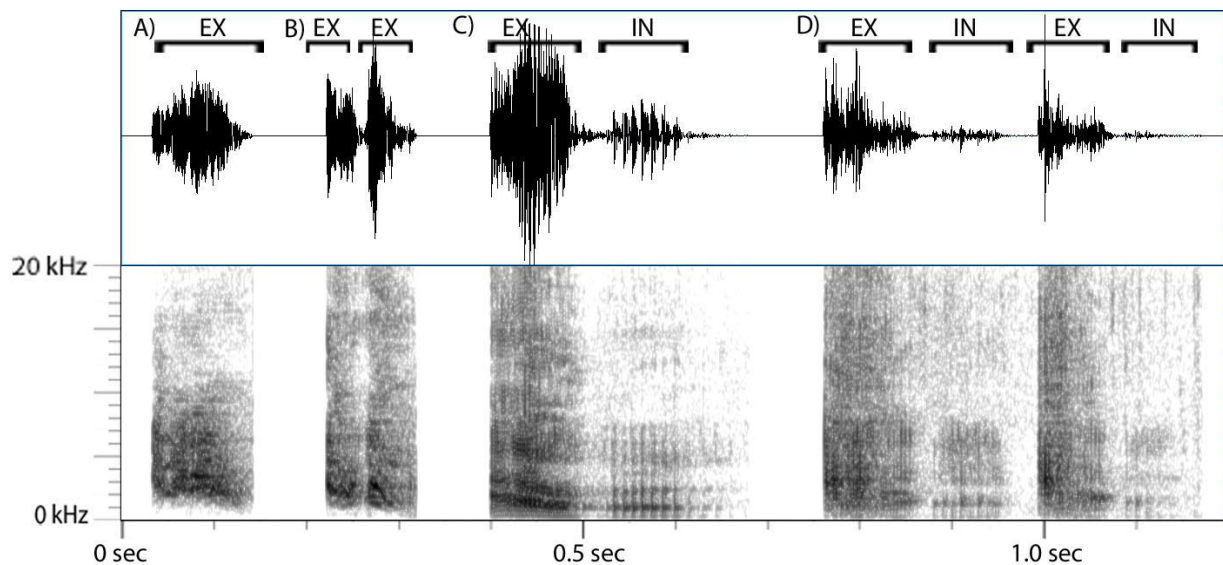


Figure 3.1: Oscillograms and spectrograms representing a) Typical alarm call produced by a female to a terrestrial predator containing a single exhaled element, B) Typical alarm call produced by a female to a terrestrial predator containing two exhaled elements, C) Typical alarm call produced by a male to a terrestrial predator containing one exhaled, and one inhaled element D) Typical alarm call produced by a male to a terrestrial predator containing two exhaled, and two inhaled elements. In each call, exhaled elements are marked with an “EX”, and Inhale elements are indicated with an “IN”. The frequency range on the spectrograms is from 0Hz to 20kHz. The entire segment is 1.2 seconds long.

Acoustic Analysis

Before analysis, calls were high pass filtered at 100Hz to remove any noise below the lowest frequency of the call (Raven Pro, v1.5). I labeled each call element, and saved them as separate files. Call elements were padded with 0.2 second silent margins using Avisoft SASLab Pro, before being resampled to 16kHz. I performed a Fast Fourier transform (Hamming window function, 1024 points, 93.75% overlap) to generate a spectrogram for each call element. The spectrograms were then exported and analyzed using a custom sound analysis software developed by K. Hammerschmidt (LMA - Lautmusteranalyse v. 2018_0.4).

It was often the case that alarm calls lacked any clear harmonic structure, making measurements of fundamental frequency impossible in many instances. Instead, I followed the procedures outlined by Price et al. (2015), and measured variables relating to the broader distribution of energy within the call. Descriptions of the acoustic variables used in this analysis are described in **Chapter 2**. A list of the acoustic variables used in this analysis are listed in Table 3.1.

Table 3.1: Acoustic parameters used in random forest models

Temporal measurements – Based on entire call	
Call Duration	Full duration of the call
Number of elements	Number of call elements within a call
Mean element length	Mean element length within a call
Spectral measurements – Based on single elements	
Pfmean	Mean peak frequency [Hz]
Pfmaloc	Location of the maximum peak frequency [(1/duration)*location]
Pfmiloc	location of the minimum PF [(1/duration)*location]
Pfjump	maximum difference between successive PF's [Hz]
Pftrfak	factor of linear trend of PF (global modulation: -1 to 1)
Pftrmean	mean deviation between PF and linear trend [Hz]
Pftrmax	maximum deviation between PF and linear trend [Hz]
DFA1	Distribution of the frequency amplitudes 1: Mean first quartile of amplitude in spectrum
DFA2	Distribution of the frequency amplitudes 2: Mean second quartile of amplitude in spectrum
DFA3	Distribution of the frequency amplitudes 3: Mean third quartile of amplitude in spectrum
mean_noise	Mean value of noise within call. 0 = pure tone, 1 = random noise
DF1	Mean frequency of 1 st dominant frequency [Hz]
df1maxloc	Location of the maximum frequency 1 st DF [(1/duration)*location]
df1minloc	Location of the minimum frequency 1 st DF [(1/duration)*location]
df1trfac	factor of linear trend of 1st DF (global modulation)

Statistical analysis

Random forest models

I used a random forest model (Breiman 2001) to, a) assess how well terrestrial alarm calls produced by males could be distinguished from those produced by females, b) determine whether alarm calls produced by different individuals could be distinguished from one another, and c) to determine which acoustic variables were best suited for distinguishing between the calls of males and females and different individuals. A detailed description of random forest analysis is provided in **Chapter 2**. Random forest models are a machine learning technique that extend and improve on standard CART (Classification And Regression Tree) analysis (Kassambara 2017a). The random forest model classifies each case in the sample to either the correct group (correct assignment), or to an incorrect group (incorrect assignment). Models were “trained” using a randomly selected 70% of the data (training data set). The remaining 30% of the data were used as the “validation data set” to determine whether the random forest models generalize well to data that were not used to train the model (Kassambara 2017a). The first random forest model was used to determine whether the terrestrial alarm calls of adult males were quantitatively different than those produced by other age/sex classes. Here, I asked the model to distinguish between the calls of adult males and females (lumped into one category). A second random forest model was used to determine how well the calls of different individuals could be distinguished based on their acoustic structure alone.

I re-analyzed a subset of the data using a series of permuted discriminant function analyses (pDFA), which allowed me to control for the influence of non-independent data within my data set (Mundry and Sommer 2007). For these analyses, I used a nested design, in which individuals are randomized repeatedly (permuted) between different age/sex classes. This

randomization process should effectively mask any discriminability between the different classes if it exists. The discriminability of each sex class in the permuted data sets is compared with the discriminability of the original data set, with the null hypothesis that the discriminant functions analysis is no better at discriminating between the true (non-randomized) groups than it is at discriminating between two permuted (random) groups. I ran all pDFAs described below using 1000 permutations, with a set of uncorrelated acoustic variables.

I also performed a series of nonparametric tests (Mann and Whitney 1947, Kruskal and Wallis 1952; Jaccard et al. 1984) to determine which of the measured acoustic variables differed significantly between adult males and females, or between adult males and all other age/sex classes.

Cluster analysis:

While the random forest model allowed me to determine how the alarms of males and females differed in structure, I performed a fuzzy, or “soft” cluster analysis to determine whether the calls of males and females fall into the distinct categories, corresponding to the ‘chirps’ and ‘threat alarm barks’ as described by Struhsaker (1967a), or whether terrestrial alarm calls are characterized by intermediate forms. This analysis also allowed me to determine whether “chirps” and “alarms” are used predominantly by females and males respectively, or whether both sexes produce each call type. The fuzzy cluster analysis was performed using the “fanny” function from the “cluster” package in R (Maechler et al. 2018). Fuzzy cluster analysis differs from classical clustering approaches in that each case (each alarm call in the data set) is assigned a probability ranging from 0 to 1 of belonging to each cluster (Kassambara 2017b). Data points that are most

indicative of each cluster (i.e., are located near the center of a cluster) are given a value closer to 1 for that cluster, and are therefore given probabilities closer to zero for other clusters. A point situated exactly between two clusters is assigned a probability of 0.5 for each cluster. Fuzzy cluster analysis thus provides a quantifiable measurement of how much a given call belongs to a given cluster, and as such allowed me to identify calls that exist as intermediates between each of the purported call “types”.

Because cluster analyses can be skewed by highly correlated variables (Kassambara 2017b), I used a set of 13 uncorrelated acoustic variables for this analysis. When two or more variables were highly correlated, I retained the variable that was rated highest in terms of its importance for discriminating between the calls of males and females by the random forest model. Importantly, while I kept in the most “important” variables from the random forest models, I also retained a number of uncorrelated variables that were rated as being less important for separating the calls of male and females. This was done to avoid pre-biasing the cluster analyses into identifying clusters that corresponded to sex differences in alarm structure. Because cluster analysis is highly sensitive to outliers, I used the Box-Cox function from the R package "MASS" (Venables and Ripley 2002) to find the optimal power coefficient with which to transform each variable to reduce skewness. The transformation used for each variable is listed in Table 3.2. The transformed variables were then standardized to make the variables comparable (Kassambara 2017b).

Table 3.2: Transformations applied to the thirteen uncorrelated acoustic parameters before scaling data for cluster analysis

Parameter	Transformation
DF1	$^{0.1}$
Number_of_elements	$^{-0.2}$
Average_element_length	$^{0.3}$
Pftrfak	$^{1.6}$
pftmean	$^{0.2}$
mean_noise	$^{0.2}$
df1trfac	$^{2.6}$
pfmaloc	$^{-0.3}$
pfmiloc	$^{1.1}$
df1minloc	no change
df1maxloc	Log
Call Duration	$^{-0.6}$
DFA3	$^{1.2}$

Determining optimal number of clusters, and evaluating the cluster solution

The fuzzy clustering algorithm requires the desired/predicted number of clusters (k) to be specified by the user. Although previous work on this species suggests that a 2-cluster solution would be optimal (A separate cluster for ‘chirps’ and ‘threat alarm barks’, respectively), I wanted to let the data dictate the optimum number of groups, allowing for an objective clustering solution based solely on the actual data. To this end, I calculated the average silhouette value for various numbers of clusters using the `fviz_nbclust` function from the “NbClust” package in R (Charrad et al. 2014). Silhouette values are a commonly used method of evaluating cluster solutions that take into account how close a case is to the center of its own cluster, and how far away it is from the center of any other neighboring cluster(s). Each case is given a value from -1 to 1, with higher values representing well clustered cases characterized by small within-cluster distance and high between-cluster distance (Kaufman and Rousseeuw 2009, Kassambara 2017). Silhouette values close to zero suggest a case is relatively distant from the center of its own cluster. Negative cluster

values represent poorly clustered cases, and could suggest that these calls belong to another cluster. The average silhouette value for the entire cluster solution can be used as a metric for how well the objects fit into their respective clusters on average. The number of clusters that maximizes the mean silhouette value is therefore considered the optimum solution (Kaufman and Rousseeuw 2009, Kassambara 2017b). Silhouette values also provide an important means by which the final cluster solution can be evaluated. Cluster solutions with average silhouette values of 0.51 and above are generally considered to represent cases where data can be partitioned into reasonably discrete clusters (Kaufman and Rousseeuw 2009). Values below this suggest intergradation between clusters.

3.4 Results

Alarm calls – General description

As described by Struhsaker (1967a), the alarm calls produced by adult males to terrestrial predators and the predator mounts were typically loud, low frequency vocalizations, with abrupt onsets. Although some calls consisted of only single units, calls in my sample contained as many as 10 elements separated by periods of acoustic energy approximating that of background noise. When calls were made up of multiple elements, they generally consisted of an exhaled element followed by alternating inhale and exhale units (Fig. 3.1C, D). A call could end on either an inhale or an exhale. Calls produced by females generally consisted of one or two closely spaced, short elements (Fig. 3.1A, B). Still, females also produced calls that contained audible inhaled units on occasion, although these units often seemed to be much lower in amplitude compared to the inhales produced by adult males.

Acoustic differences between male and female alarm calls

I used a random forest model to determine how well the alarms of males and females could be classified based on acoustic structure alone. The model was able to distinguish between the calls of males and females with a high rate of accuracy (98.01% accuracy, OOB error = 1.99%). Classification error was low for calls produced by both males, and females (Males = 0.01, Females = 0.05). The model accurately predicted the caller's sex 96.47% of the time from the validation data set (95% CIs: 90.03, 99.27), which was significantly better than chance expectation ($p < 0.001$, expected classification accuracy by chance = 60.58%). Calls produced by both males and females were classified correctly at levels that exceeded chance (Adult males: 96.77 % accuracy vs 72.98 % accuracy expected by chance, $p < 0.001$; Adult females: 95.65 % accuracy vs 27.16 % accuracy expected by chance, $p < 0.001$).

A permuted discriminant functions analysis which controlled for caller identity was able to correctly classify 78.35% of the calls correctly, which was significantly better than the 52.12% expected accuracy by chance ($p = 0.001$). These results suggest that the calls produced by adult males can be clearly discriminated from those of adult females.

Table 3.3: Confusion matrix: The number of correct classification and misattributions for each sex (adult males vs adult female). Correct classifications are highlighted in green. Misclassification rate represents the percent of calls ascribed to each caller category that were misclassifications. For each caller category, permutation tests were used to compare the percent of calls correctly classified to the percent expected by chance.

		Actual caller category		Misclassification rate
		Adult Male	Adult Female	
Predicted Caller category	Adult Male	60	1	1.64%
	Adult Female	2	22	8.33%
Classification accuracy		96.77%	95.65%	
Expected accuracy by chance		72.98 %	27.16 %	
P-value		<0.001	<0.001	

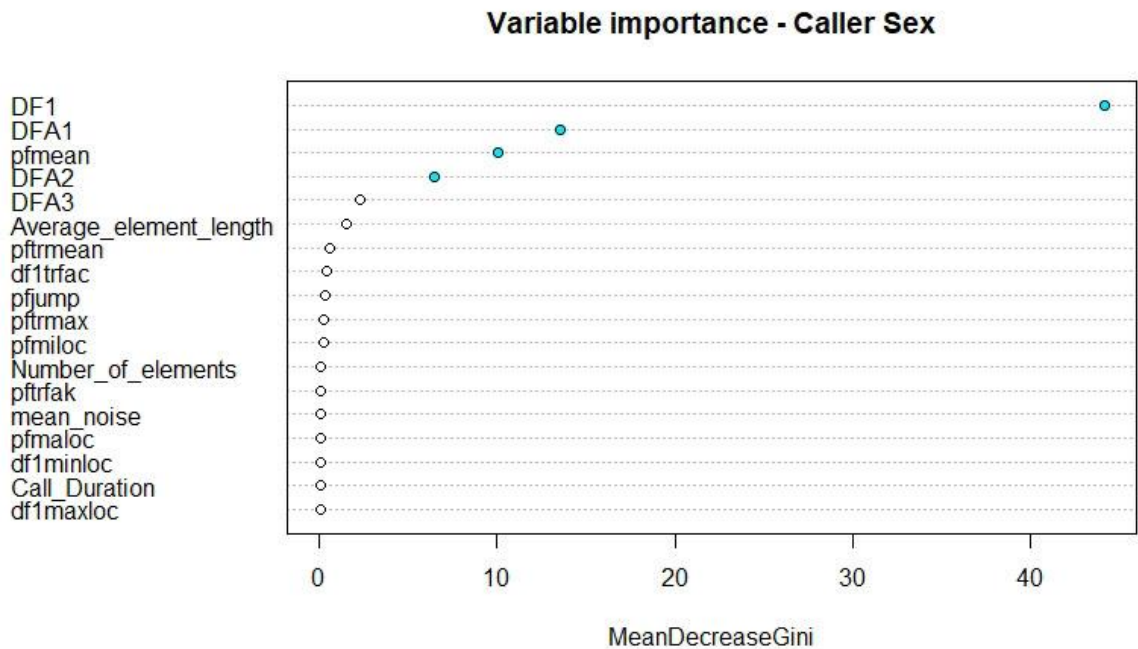


Figure 3.2: Variable importance plot for predictor variables from random forest classifications used for predicting caller sex. The “mean decrease in Gini” index is the average decrease in node impurity that results from splits using a given acoustic variable. Higher values of mean decrease in Gini indicate variables that are more important to the classification. Cyan dots represent Gini indices higher than the mean Gini index for all variables in the analysis. Descriptions of each acoustic variable are provided in Table 3.1.

Of the 20 acoustic variables entered into the model, four were rated as having higher Gini values than the average Gini index for all variables (average Gini index = 4.47); DF1, DFA1, peak frequency, and DFA2 (Fig. 3.2).

Of these four variables, DF1 was clearly more important for distinguishing between caller sex than all others. A series of Wilcoxon rank sum tests showed that all four variables differed significantly between the calls produced by adult males and adult females (Fig. 3.3); compared to females, the alarm calls produced by males had a lower first dominant frequency band ($W = 786.5$,

$p < 0.001$), lower values for DFA1 ($W = 1264.5$, $p < 0.001$), lower peak frequencies ($W = 1444.5$, $p < 0.001$), and lower values for DFA2 ($W = 1546.5$, $p < 0.001$).

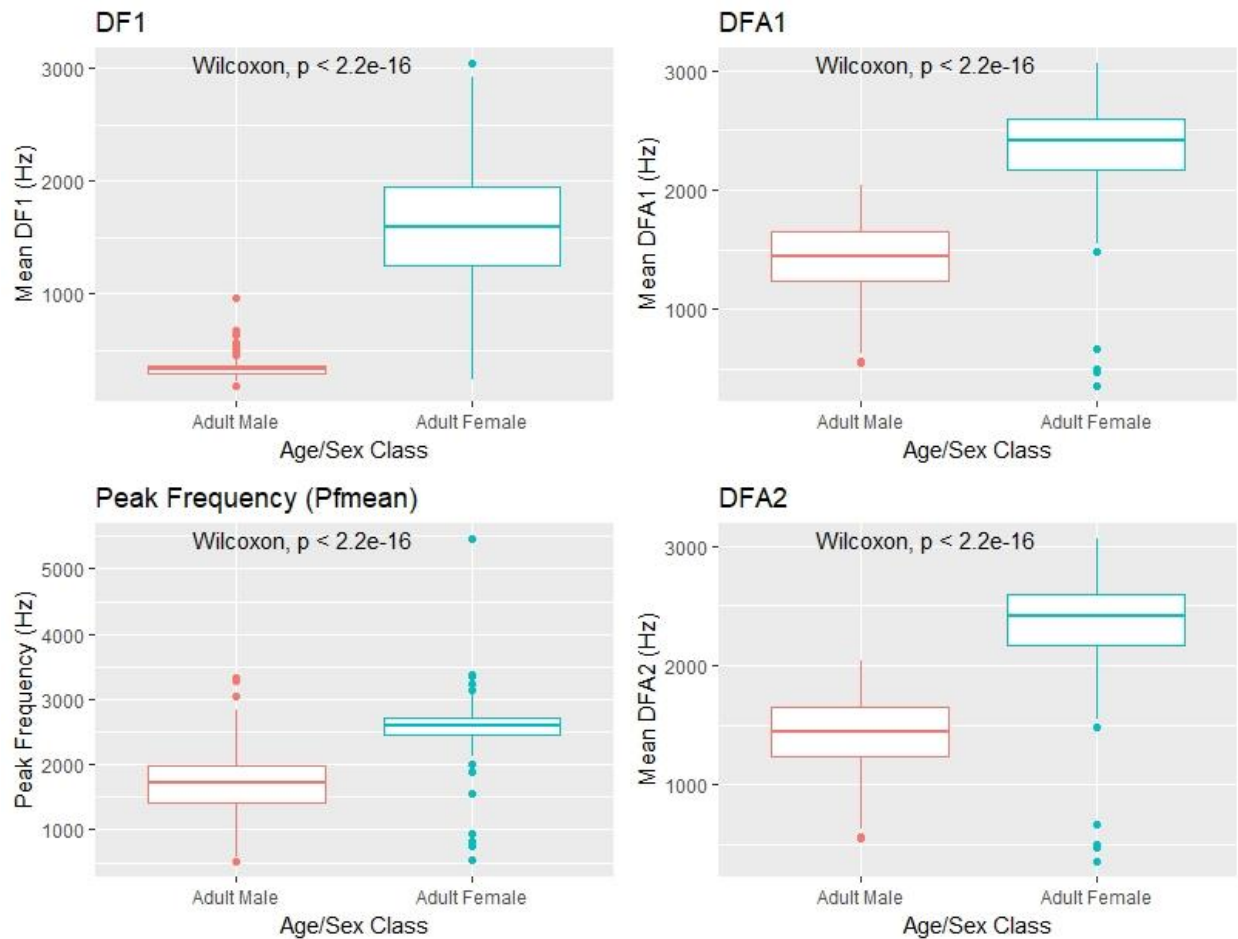


Figure 3.3: Boxplots representing the distribution of four variables which vary significantly between adult males and adult females; DF1, DFA1, pfmean, and DFA2. In each box plot, the center line represents the median. The lower and upper hinges (bottom and top of each box) represent the first and third quartiles of the data respectively. Upper whiskers extend to the largest value that is no further than 1.5 times the inter quartile range (distance between the first and third quartiles) from the upper hinge. The lower whiskers extend from the lower hinge to the smallest value that is no more than 1.5 times the inter quartile range. Points outside the whiskers represent outliers.

Cluster analysis:

Although the results of the random forest model suggest that the calls of males can be distinguished from those produced by females to a high degree of accuracy based on their acoustic structure alone, I wanted to determine whether these alarm calls exist as distinct call “types”, or whether the alarm calls of males and females are characterized by intergradation in acoustic structure.

The silhouette analysis revealed that 2-clusters would provide the optimum solution for the adult data set (Fig. 3.4). The two-cluster solution resulted in clusters containing 187 and 99 calls respectively (Table 3.4). While both clusters contained calls produced by both males and females, 87.44% of calls produced by adult males fell into cluster 1. Conversely, 92.41% of calls produced by adult females fell into cluster 2. A chi-square test (MASS package, `chisq.test` function) revealed that call category was not independent of caller sex (161.057, $df = 1$, $p < 0.001$), with adult males falling into the 1st cluster at higher rates than expected by chance, and adult females falling into the 2nd cluster at rates higher than expected by chance.

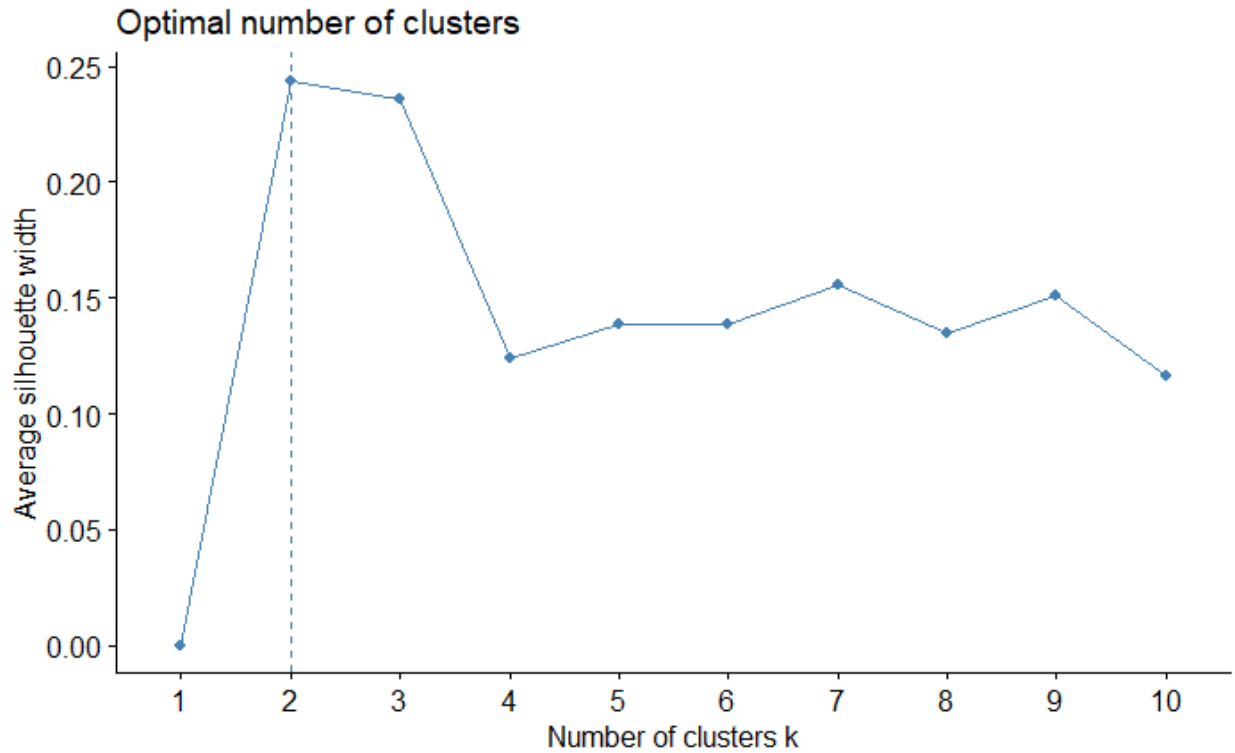


Figure 3.4: Silhouette values for the cluster analysis as a function of the number of clusters. Cluster analysis is based on thirteen uncorrelated acoustic variables.

Table 3.4: Distribution of males and females between the two clusters

	Male	Female
Cluster 1	181	6
Cluster 2	26	73

The mean silhouette value for the entire solution was 0.24. Cluster 1 (the primarily male cluster) had a silhouette value of 0.36, while the second cluster (primarily female cluster) had a silhouette value of 0.02, thus suggesting that relative to cluster 1, the calls in cluster 2 were less stereotyped in structure. Still, the silhouette value of 0.36 does not suggest clearly structured data

(Kaufman and Rousseeuw, 2009), suggesting that a number of calls within this cluster differed substantially in structure from calls near the center of the cluster (Fig. 3.5).

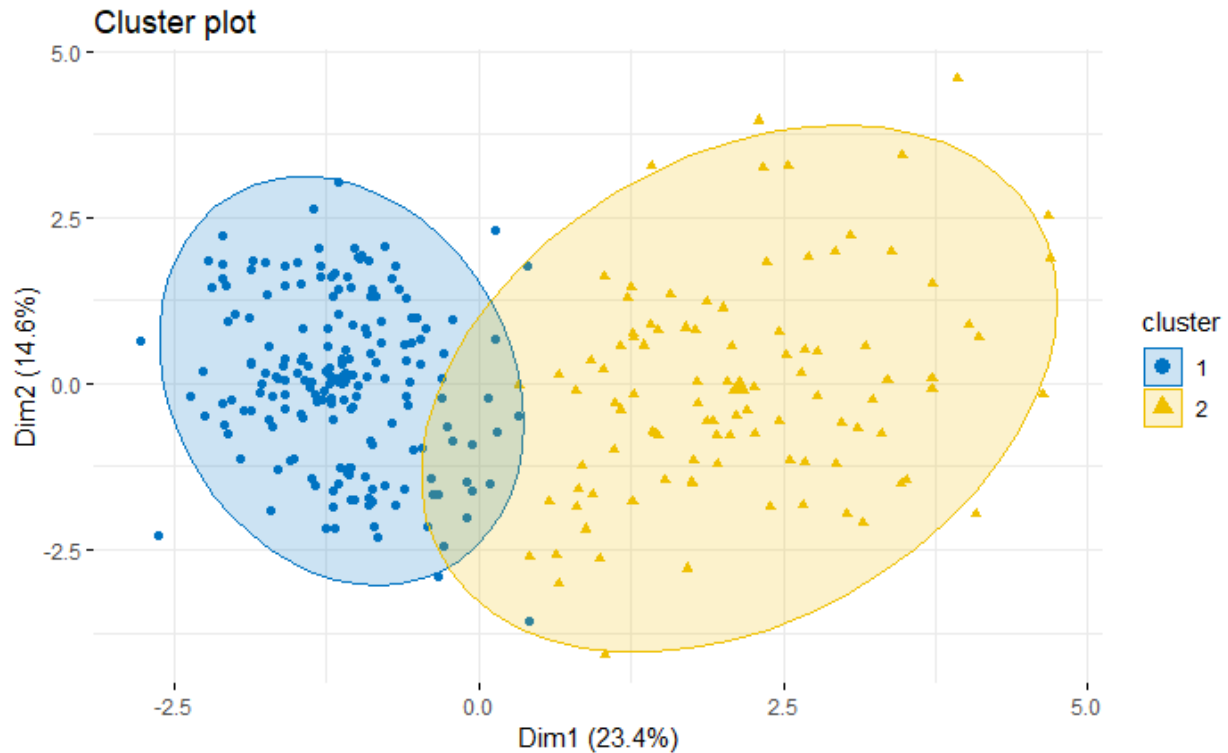


Figure 3.5: Results of the soft cluster analysis: Scatterplot represents the calls of adult males and adult females plotted against the first two principal components. Blue dots represent calls in cluster 1, which was comprised principally of calls produced by adult males. Yellow triangles represent calls in cluster 2, which was comprised principally of calls produced by adult females. Principal component 1 (23.4% of variance) represents element length (0.823), DFA3 (-0.774), df1minloc (0.732), DF1 (-0.686), Mean noise (-0.599). Principal component 2 (14.6% of variance) represents Number of elements (0.851), Full call duration (0.763), and df1trfac (0.523).

The membership values for calls produced by males and females are represented in figure 3.6. While calls produced by males and females generally fell into separate clusters, it was clear from the membership values that males did produce calls on occasion that were closer in structure

to the stereotypical calls produced by females. Additionally, both males and females seemed to produce calls on occasion that fell between the two clusters, suggesting some intergradation between call types (Fig. 3.7).

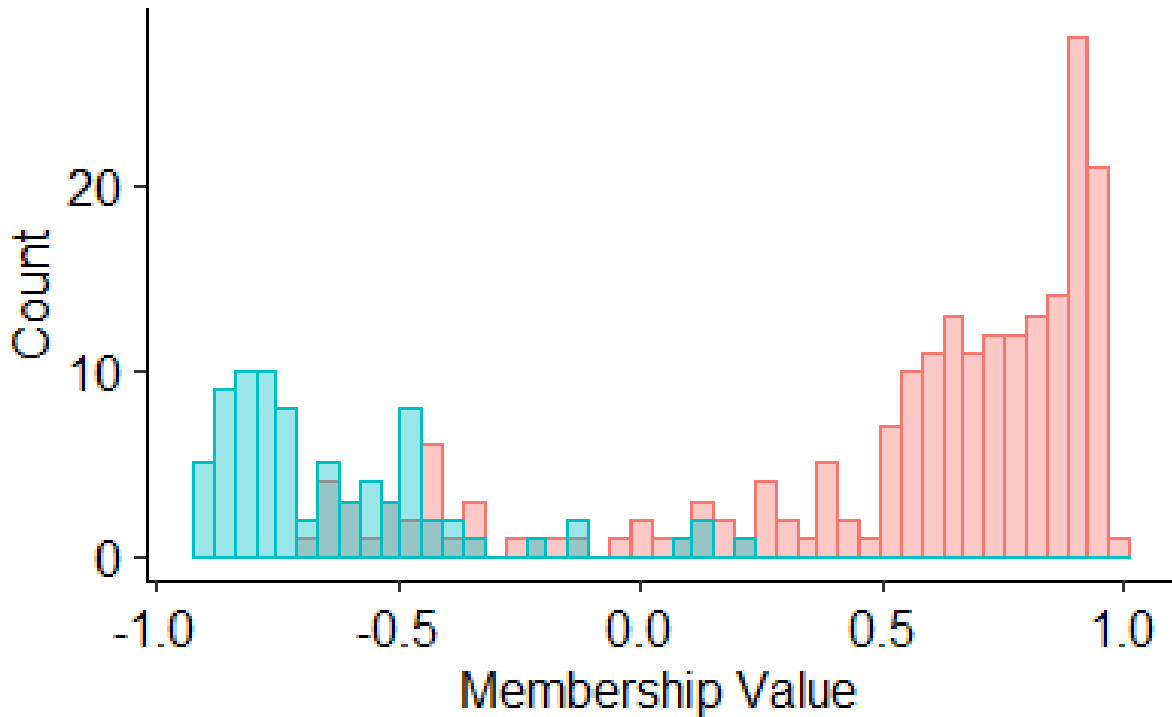


Figure 3.6: : Histogram representing cluster membership. Red bars represent calls produced by males, and Blue bars represent calls produced by females. Membership value is calculated as the membership value of a call for cluster 1 minus its membership value for cluster 2. Membership values of 1 and -1 suggest a call is typical of either cluster 1 or cluster 2 respectively. Call with membership values closer to 0 represent intermediates relative to the two clusters.

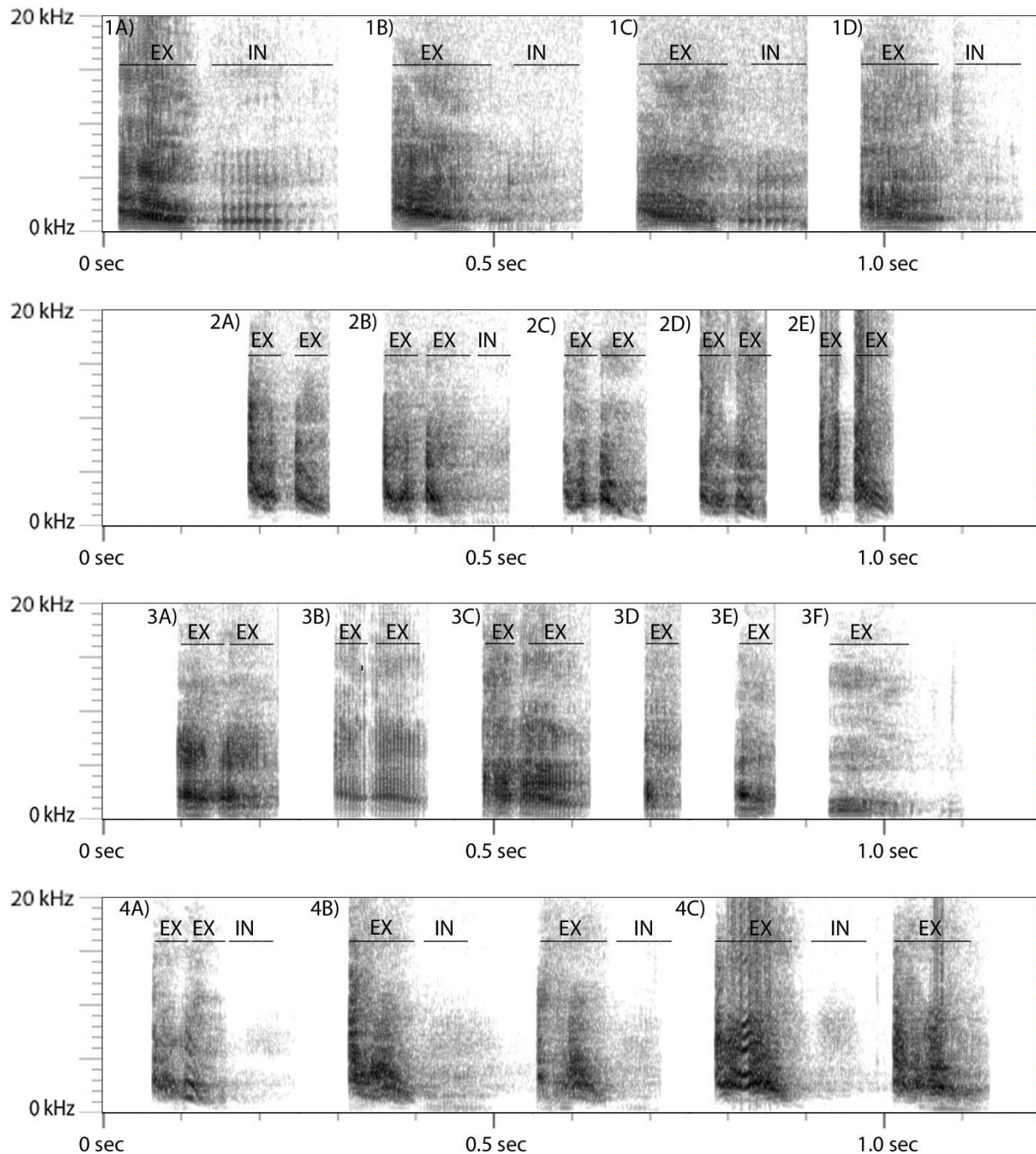


Figure 3.7: Call 1A, 1B, 1C, and 1D represent alarm calls that were typical of the first cluster. All four calls are produced by adult males. Calls 2A, 2B, 2C, 2D, and 2E represent calls that were typical of the second cluster. All five calls are produced by females. Calls 3A, 3B, 3C, 3D, 3E, and 3F represent calls produced by males that had more intermediate membership values for both clusters. Calls 4A, 4B, and 4C represent calls produced by females that were more intermediate between the two clusters. Portions of the call labeled “EX” represent exhaled portions of the call, while portions labelled “IN” represent audible inhaled portions.

Classification by caller ID: Individual differences in alarm calls

I ran a second random forest model to determine whether the alarm calls produced by different individuals were discriminable from one another. The model was able to distinguish among calls produced by different individuals at a relatively high rate (66.02% accuracy, OOB error = 33.98%). When looking at the validation set, the model accurately predicted the identity of the caller 69.84% of the time, which was better than chance expectation (Expected accuracy by chance = 13.66%, $p \leq 0.001$). Although all individuals were classified at levels that were better than chance expectation (Table 3.5), the calls of some individuals were classified incorrectly at relatively high rates, suggesting that some individuals may produce calls that are more clearly individually recognizable than others. Variable importance for discriminating between the calls of different individuals are illustrated in Fig. 3.8.

Table 3.5: Confusion matrix - The number of correct classification and misattributions for each individual. Correct classifications are highlighted in green. Misclassification rate represents the percent of calls ascribed to each caller (predicted caller) that were not actually produced by that caller. For each individual, permutation tests were used to compare the percent of calls correctly classified to the percent expected by chance.

		Actual Caller								Misclassification Rate
		AM1	AM2	AM3	AM4	AF1	AF2	AF3	AF4	
Predicted Caller	AM1	4	1	0	0	0	0	0	0	20.00%
	AM2	0	9	0	2	0	0	0	0	18.18%
	AM3	1	0	12	2	0	0	0	0	20.00%
	AM4	3	1	0	4	0	0	0	0	50.00%
	AF1	0	0	0	0	3	0	0	0	0.00%
	AF2	0	0	0	0	0	2	0	0	0.00%
	AF3	0	0	0	0	1	0	5	2	37.50%
	AF4	0	0	0	0	0	4	2	5	54.55%
Percent correctly classified		50.00%	81.82%	100.00%	50.00%	75.00%	33.33%	71.43%	71.43%	
Expected correct classification by chance		12.78%	17.43%	19.08%	12.44%	6.23%	9.71%	11.08%	11.02%	
P-Value		0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	

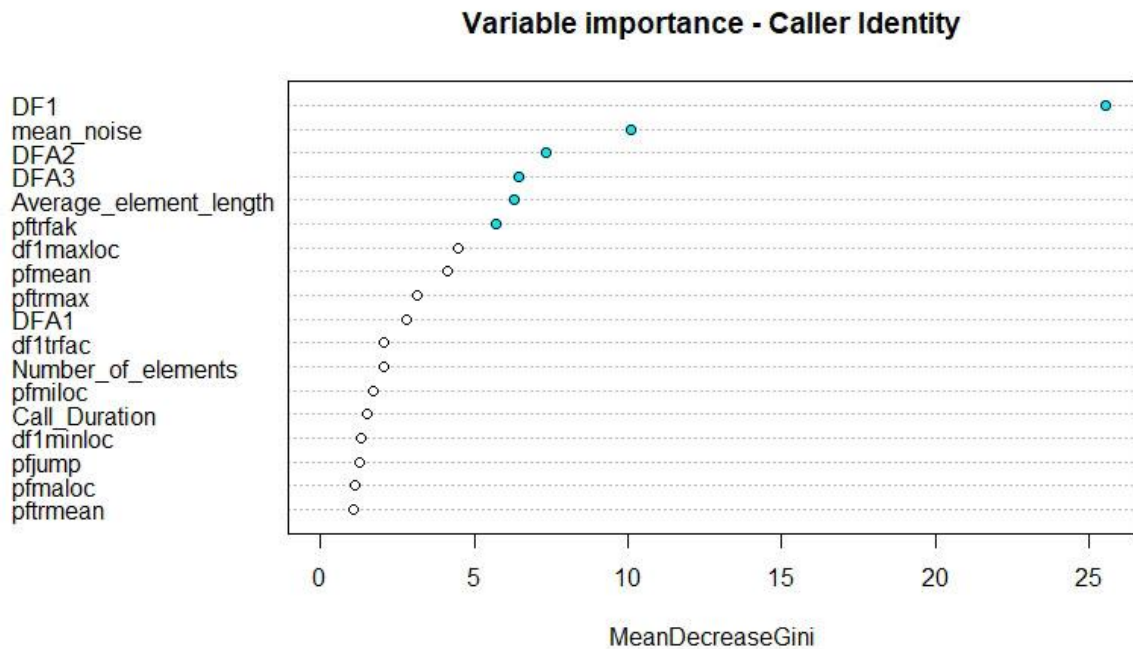


Figure 3.8: Variable importance plot for predictor variables from random forests classifications used for predicting caller identity. The “mean decrease in Gini” index is the average decrease in node impurity that results from splits using a given acoustic variable. Higher values of mean decrease in Gini indicate variables that are more important to the classification. Cyan dots represent Gini indices higher than the mean Gini score for all variables in the analysis. Descriptions of each acoustic variable are provided in Table 3.1.

3.5 Discussion

In this chapter, my goals were to determine how well the alarm calls produced by adult males and females could be distinguished from one another based on their acoustic structure, and to determine which acoustic parameters most clearly separate alarm calls by caller sex. I also attempted to determine whether the terrestrial alarm calls of adult male and female vervet monkeys exist as two distinct call ‘types’, or whether “chirps” and “threat alarm barks” are characterized by acoustic integration between call types. Finally, I investigated whether alarm calls vary in structure between different individuals.

My analyses revealed that the alarm calls produced by adult males are quantitatively distinguishable from those produced by adult females. Generally speaking, males produced alarm calls that had lower first dominant frequency bands, lower peak frequencies, and lower overall energy distribution relative to females. One hypothesis to explain the observed sex differences in the structure of terrestrial alarm calls is that call structure simply reflects the broad differences in the structure of the male's vocal apparatus relative to females. Vervet monkeys are sexually dimorphic at adulthood, with males being larger and weighing more than females (Turner et al. 1997 - females average 3.3kg, and males 5.5kg at the study site; Pasternak et al. 2013). The observed sex differences in call structure are in line with what would be expected based on the fact that adult males are larger than all other age sex classes. Sex differences in body size are expected to impact the frequency content of a signal, with larger vocal folds and longer pharynges producing vocal signals with increased energy content in the lower frequency ranges (Fitch and Hauser 2003, Ey et al. 2007a). Additionally, the length and cross-sectional area of the vocal tract, which varies with the size of a caller, affects the frequencies that are attenuated or amplified during phonation. Vocal tract length is inversely proportional to the spacing between frequency peaks within a given signal, and as such, smaller individuals (i.e., females) are expected to have frequency peaks that are distributed more widely than larger (adult males) individuals (e.g., *Macaca mulatta*; Fitch 1997; *Canis familiaris*; Riede and Fitch 1999; *Homo sapiens*; Fitch and Giedd; 1999). As such, it is of course possible that the tendency for males to produce a call that is acoustically "different" than females is simply a by-product of sex differences in body size at adulthood.

Still, anatomical constraints cannot explain why adult males did on occasion produce vocalizations that grouped strongly into the predominantly female cluster (meaning they were actually more acoustically similar to the "chirp" generally produced by females). If the calls'

acoustic structure is limited by the characteristics of the vocal apparatus, males should not have been able to produce signals that are acoustically similar to those generally produced by females. Thus, another, non-mutually exclusive possibility is that sex differences in call structure represent a sexually selected trait that is not just a by-product of sexual selection on body size. Because the alarm calls produced by males seem to diverge in structure around puberty (Struhsaker 1967a), they in some ways mirror the developmental patterns of other sexually selected traits (Darwin 1896, Kappler and van Schaik 2004). Similar interpretations have been put forward for the sexually distinct alarm calls of Diana monkeys (*Cercopithecus diana*), which are also characterized by sexual dimorphism at adulthood (Zuberbühler et al. 1997). The vocal exaggeration hypothesis suggests that in some instances, it may prove beneficial for signalers to produce vocalizations that are characterized by acoustic parameters that acoustically exaggerate their body size (Fitch 1999). Still, to investigate this possibility further, the simple effects of sex differences in body size must be ruled out as an explanation.

Intergradation between call types

Despite the apparent age/sex differences in call structure, my cluster analysis revealed that the alarm calls produced by males and females do not exist as two distinct call ‘types’, as has been suggested by prior research (Struhsaker 1967a). Although alarm calls did tend to group into two clusters that were largely composed of calls produced by adult males and adult females respectively, these clusters were characterized by at least some acoustic overlap, as highlighted by the relatively low silhouette values for the 2-cluster solution. Additionally, while some calls possessed membership values close to 1 for their given cluster, others possessed membership

values closer to 0.5, suggesting they exist as acoustic intermediates between the more stereotypical calls that represented each cluster.

Intergradation in call structure within and between contexts is often interpreted as evidence that acoustic variation corresponds to graded changes in the internal, affective states of signalers (Solitis et al. 2005, Coss et al. 2007, Dunlop 2017; Darwin 1872, Hammerschmidt and Fischer 1998). One possibility then is that the graded variation that characterizes the acoustic structure of terrestrial alarm calls mirrors variations in the internal state of different callers within the context of a predator encounter. During a bout of alarm calling, it was my impression that males often moved into trees that were closer to the eliciting predator or predator mounts. Conversely, females and immature individuals generally maintained their distance. These differences in behaviour suggest that the sexes may vary in their motivational states during predator encounters, which could be reflected in the structure of the alarm calls that they produce. In line with this hypothesis, Seyfarth and Cheney (1980) also noted that females were more likely to produce alarm calls that were subjectively more similar to those produced by males when they were closer to the predator that elicited the signal.

It has been suggested that the vocal signals produced by non-human animals are largely mediated by emotion and motivation (Darwin 1872, Morton 1977). Importantly, signals that “refer” to specific external contexts are still likely to vary in structure based on the internal state of callers (Marler et al. 1992, Manser et al. 2002, Seyfarth and Cheney 2003). For example, as argued for vervet monkeys, suricates (*Suricata suricatta*) produce three alarm calls to different predator classes that elicit different, adaptive responses in receivers (Manser et al. 2002). Still, within this system, alarm calls vary in structure based on the perceived threat of a given predator encounter, suggesting that a caller’s emotional state still affects the call’s structure within a given

context. Importantly, receiver responses to calls produced within predator categories vary in intensity based on this motivational variation. While the alarm calls of various taxa were once categorized as being either referentially *or* motivationally based, it is now generally accepted that this dichotomy is an over simplification (Manser et al 2002, Seyfarth and Cheney 2003; Fischer 2017). So-called referential signals are likely to contain acoustic cues to a caller's motivation, in much the same way human words are both semantic and express emotion while speaking (Seyfarth and Cheney 2003). While I do not suggest here that vervet alarm calls are "word-like", it is likely that context-specific vocalizations still vary a great deal in structure depending on a caller's internal motivation at the time of calling. Following this, the "chirp-like" and "bark-like" calls produced by vervet monkeys upon encountering a terrestrial predator might simply represent variations in a caller's internal state. This "caller urgency" hypothesis could be tested, for instance, by experimentally altering the size of predator mounts, or the distance between mounts and callers (see Warkentin et al. 2001, Manser 2001), in an attempt to determine whether the perceived threat of the eliciting stimulus has an effect on whether callers produce chirp-like or bark-like alarms.

Individual distinctiveness

My analysis revealed that vervet terrestrial alarm calls contained cues regarding caller identity. Many of the most important variables for individual discrimination relate to the frequency content of the calls (e.g., DF1, DFA2, DFA3). These variables are likely to be dependant on the properties of a caller's vocal tract or vocal folds, which are likely to vary in structure among individuals. Similarly, inter-individual differences in the size of a caller's lungs may explain why mean element length was rated as being important for individual discrimination of calls as well;

variation in caller body size or lung volume would affect the length of vocalizations produced by signalers.

Perhaps the most interesting finding here is that vervet monkey alarm calls are more individually recognizable than vervet grunts, which I found did not exhibit quantifiable, statistically significant differences among individuals (see chapter 2). Asymmetries in the individual distinctiveness of different call types within a species vocal repertoire have been observed in other species as well (Rendall et al. 2009b, Charrier et al. 2001). Mechanistically, the acoustic structure of different call types is likely to influence the degree to which they can reflect the structure of a caller's vocal tract. For instance, vocal signals that have low fundamental frequencies, and that are rich in harmonics are more likely to convey cues related to vocal tract structure than calls that are less harmonically rich, such as pure tone signals, or even more harmonically rich signals produced at low amplitudes (Owren and Rendall 2001). Functionally, it is possible that not all calls are under the same selective pressure to contain cues to a caller's identity. It is interesting, however, that in this instance, my results indicate the opposite pattern to what might be expected based on the contexts that seem to elicit calling. If the only function of an alarm call is to warn conspecifics as to the presence of a potential threat, the identity of the caller should be irrelevant. Conversely, grunts, which are produced in social contexts, may be expected to convey information about caller identity, which plays a deterministic role in many social interactions. Still, as explained elsewhere (**Chapter 2**), grunts are generally given in such close proximity, that cues to caller identity may be unnecessary. However, this does not explain the individual discriminability of the terrestrial alarm call.

One possibility is that individual cues to caller identity allow receivers to substantiate the presence of an actual threat, based on perceiving signals from multiple sources. Playback

experiments with cotton top tamarins (*Saguinus oedipus*) revealed that subjects habituated to alarm calls produced by a single individual, but dishabituated on hearing the alarms produced by another individual (Sproul et al. 2006), suggesting that, from the perspective of a receiver, the alarm calls produced by a second individual may affirm the continued presence of an actual threat. A related possibility is that vervet monkeys may make use of individual specific cues to caller identity to assess caller reliability. Evidence that receivers take caller reliability into account when responding to vocal signals has been demonstrated in a number of species (e.g., vervet monkeys: Cheney and Seyfarth 1988; bonnet macaques, *Macaca radiata*: Ramakrishnan and Coss 2000; California ground squirrels, *Spermophilus beecheyi*: Hanson and Coss 2001; Richardson's ground squirrels, *Spermophilus richardsonii*, Hare and Atkins 2001; marmots, *Marmota flaviventris*: Blumstein et al 2004). For example, Cheney and Seyfarth (1988) found that upon hearing repeated playbacks and habituating to the intergroup 'wrr' vocalizations of a given individual in the absence of another social group (simulating a false alarm, and an unreliable signaler), vervet monkey's transferred this habituation to the playback of an intergroup 'chutter' (which has the same purported referent; the detection of another social group) from the same individual. Conversely, when played the intergroup chutter of a different individual after the habituation process, vervets did not transfer this habituation, suggesting they take the identity and reliability of callers into account when responding to the calls of conspecifics. It is possible then that vervet monkeys could use cues to caller identity to discriminate between the alarm calls of different individuals based on their past reliability. Another possibility is that vervet monkeys may use individual cues to caller identity to assess the number of individuals who are calling during a given predator encounter. Sloan and Hare (2008) suggest that adult Richardson's ground squirrels may assess the urgency of a given predator encounter by attending to the number of animals producing alarm calls. They found that

adults were more vigilant, and maintained vigilance for longer, after hearing playbacks of two individuals producing alarm calls compared to bouts comprising a single caller. Because higher urgency situations generally elicit calling from a greater number of individuals (Sloan and Hare 2008), one possibility is that ground squirrels use their ability to discriminate between the calls of different individuals to assess the number of callers, allowing them to make inferences about the urgency of a given predator encounter (Sloan and Hare 2008). Finally, the ability to discriminate between different callers may allow receivers to track the movements of predators through the environment (Thompson and Hare 2010).

Given the likely relationship between body size, and many of the variables that seemed to vary predictably between callers, it is also possible that the observed individual variation reflects a caller's resource-holding potential or competitive ability. As previously mentioned, calls produced in response to terrestrial predators are also associated with other high-urgency situations in some species, including within/between group aggression (Fichtel and Kappeler 2002, Fischer et al. 2002, Price et al. 2015), or in contexts associated with competition over access to resources (Wheeler 2009). From a functional perspective, it is of course possible that alarm calls produced in competitive contexts might represent a form of deception, whereby 'false' alarm calls are produced as a means of driving away competitors (Wheeler 2009). Conversely, producing alarm calls in these contexts may serve some sort of competitive function more directly. As already mentioned, the acoustic properties of the male's terrestrial alarm calls seem well adapted to act as an advertisement, or even an exaggeration of their body size. As such, these calls may provide cues relating to a caller's ability, or even their motivation to compete over access to resources. Cues to caller identity may allow receivers to identify the calls of particularly competitive individuals, thus increasing the signal's functionality in competitive interactions. Evidence for this

hypothesis comes from the fact that the terrestrial alarm calls of adult males grade acoustically into signals produced during male/male aggressive interactions (Price et al. 2015).

During an encounter with a terrestrial predator, vervet monkeys produce vocalizations that have been argued elicit adaptive escape responses from receivers (but see Ducheminsky et al. 2014). The results of my analyses highlight the fact, however, that the acoustic structure of a particular alarm call may reflect specific aspects of the caller, such as its sex and identity. It remains to be determine whether cues to caller characteristics are salient to receivers and, by extension whether their salience is functional with respect to caller fitness. Future studies would benefit from focusing on responses to the different alarm call types identified here (i.e., chirps vs barks), and whether one call type elicits a stronger and/or specific response than another. In addition, given the fact that alarm calls are individually distinctive, we should look at whether callers respond differently depending on the identity of the caller. Additionally, the possibility that the sex differences in call structure that characterize the terrestrial alarm calls of vervet monkeys represent a sexually selected trait should be examined further. First, the possibility that sex differences in alarm structure is a simple by-product of sex differences in overall body size should be examined in detail. Additionally, analyses should focus on whether variation in the structure of alarm calls corresponds with caller fitness or resource holding potential (i.e., caller's dominance status, fighting ability).

Chapter 4 Vocal allometry in vervet monkeys

4.1 Introduction

Body size is likely to play an influential role in determining the outcome in a number of social situations, including aggressive interactions, competition over access to resources, and mate attraction (Clutton-Brock et al. 1979; Fitch 1999; Blanckenhorn et al. 2000; McElligott et al. 2001; Serrano-Meneses et al. 2007; Sanvito et al. 2007). As such, it is likely that animals have faced selection to make adaptive decisions based on assessments of body size and the competitive abilities of conspecifics (Given 1987, Bee et al. 1999, Masumoto 1999, Reby et al. 2005, Charlton et al. 2007a,b, Aisenberg 2009, Charrier 2011, Wyman et al. 2012, reviewed in Morton 2017). One way that animals can both advertise their own - and assess each other's - body size is through vocal communication. This is because vocalizations can act as *indices* to various signaler characteristics due to the physical relationship between a signaler's vocal production system and the calls they produce (Fitch 1999; Taylor and Reby 2010). An *index* can be defined as a signal or cue whose structure is causally linked to some relevant aspect of the signaler (Smith and Harper 2003). This relationship means that indices are unable to be faked, and as such, they allow receivers to make accurate assessments about some aspect of a signaller.

The source filter theory of vocal production has proven useful in providing a framework within which a signal's structure can be understood in terms of its relationship to the physical characteristics of a caller (Fitch and Hauser 2003; Taylor and Reby 2010). The acoustic properties of vocal signals are strongly influenced by variations in the structure of the vocal tract (Fitch and Hauser 2003, Taylor and Reby 2010) and, as such, inter-individual variation in the vocal tracts of different callers will affect the acoustic structure of the vocalizations they produce. Indices relating

to caller body size can be present in vocal signals because of the broad differences in the size and structure of their vocal tracts (Fitch and Hauser 2003; Taylor and Reby 2010). Because lung size and body size are correlated, larger individuals are able to produce louder and longer vocalizations than smaller individuals (Fitch and Hauser 2003; Taylor and Reby 2010). Similarly, because larger individuals are likely to have longer and thicker vocal folds, they also tend to produce calls with lower fundamental frequencies (Fitch and Hauser 1995, Hammerschmidt et al. 2000, Pfefferle and Fischer 2006, Ey et al. 2007a). The spectral energy of signals produced by larger animals also tends to be focused in lower frequencies compared to signals produced by smaller individuals, due to the different filter functions of larger vs smaller supra-laryngeal vocal tracts (Ey et al. 2007a). In many instances, the relationship between call structure and body size can thus be described as isometric, in that vocal parameters in theory scale proportionately with changes in body size.

Although the causal relationship between vocal tract structure and the acoustic properties of vocal signals is likely to produce honest indices to body size in many instances (Fitch and Hauser 2003), a number of authors have highlighted examples where call structure and body size have become decoupled from one another. For example, there is evidence that in some instances, selection has resulted in vocalizations that vary in acoustic structure in a way that manipulates the perceived size of callers (Morton 1977, Fitch 1999, Morton 2017). Morton (1977) highlighted the fact that in many species of birds and mammals, vocal signals produced during aggressive interactions tend to be low-frequency and noisy (perceived as harsh in tone, due to a lack of harmonic structure), and that these features are likely to be adaptive because they act as an acoustic exaggeration of a caller's body size in contexts where larger animals are likely to be at an advantage.

Researchers interested in the paralinguistic aspects of human speech have often focused on the sex differences that characterize the vocal tracts of adult humans, which exceed overall sex differences in body size (Fitch and Giedd 1999; Harries et al. 1998; Owren et al. 2007). At puberty, the male larynx increases in size, and the vocal folds both lengthen and undergo a number of structural changes which result in a considerable drop in fundamental frequency relative to females (Hollien et al. 1994, Harries et al. 1998). Additionally, males experience a secondary descent of the larynx, which has the effect of lengthening the vocal tract resulting in lower, more closely spaced formant frequencies (vocal resonances; Fitch and Giedd 1999). As such, the anatomy of the vocal tract in human males seems well adapted to produce signals that functionally exaggerate their body size relative to like-sized females. In this sense, the acoustic properties of speech in males are hyperallometric, in that at puberty, vocal pitch and formant dispersion change disproportionately relative to the speaker's body size. There is a great deal of evidence to suggest that vocal exaggeration of body size in adult males is the product of sexual selection. For instance, research has shown that men with lower voices are rated as being more physically and socially dominant by other men (Puts et al. 2006), and that females perceive voices with lower formant dispersion to be produced by men who are larger, more attractive, hairier, older, and more likely to have a muscular body type (Collins 2000). There is also evidence that men lower the pitch of their voice when talking with another man who they perceive as being less physically dominant to themselves (Puts et al. 2006). Together, these findings suggest that human male voices may face selection for the exaggeration of characteristics such as body size. Interestingly, neither fundamental frequency nor formant dispersion (the average difference between successive formant frequencies) are accurate predictors of body size among adult human males (Rendall et al. 2005), suggesting that assessment of caller body size based on these parameters in humans is likely to be

inaccurate, even within sexes. As a result, some researchers suggest that frequency content may be an index of some other aspect of individual males that receivers might use to assess quality (reviewed in Puts et al. 2014).

Similar anatomical modifications of the male vocal tract have been found in a number of non-human taxa. Male red and fallow deer (*Cervus elaphus*, *Dama dama*) possess relatively low larynges compared to most other non-human animals, and even retract their larynx while producing specific vocalizations (Fitch and Reby 2001). Both of these adaptations effectively increase the length of the vocal tract, creating the impression of a larger animal. Some authors have suggested that this allows callers to better repel males, or attract females (Fitch and Reby 2001). Tracheal elongation has been observed in a number of species of birds, and has similarly been proposed to function as an acoustic exaggeration of body size (Fitch 1999). Among primates in general, higher levels of sexual dimorphism in fundamental frequency have been found in species with polygynous mating systems, where sexual selection on males is predicted to be more intense (Puts et al. 2016). Importantly, when controlling for phylogeny and mating system, dimorphism in fundamental frequency is greater in species with lower body size dimorphism, suggesting that sexual dimorphism in call structure is not simply the by-product of differences in overall body size. Instead, it has been suggested that vocal dimorphism in these species may alleviate the need for direct, physical aggression between males (Puts et al. 2016).

Vervet alarm calls

The terrestrial alarm calls of vervet monkeys are characterised by sex differences in their acoustic structure at adulthood (Struhsaker 1967a; **Chapter 3**). While some degree of

intergradation exists between the calls of adult males and females (**Chapter 3**), they differ from one another based on their spectral profiles; relative to the “chirp” vocalizations produced most commonly by females, the acoustic energy in the male’s “threat alarm bark” is concentrated at lower frequencies. Of course, it is possible that the observed sex differences in the acoustic structure of these alarms are simply a by-product of the sex differences in overall body size that characterize this species at adulthood (Turner et al. 1997). This being the case, sex differences in the structure of alarm calls may simply reflect isometric scaling (neutral allometry) between body size and the physical structure of vocal signals in males and females respectively. However, the observed acoustic differences, when considered alongside the subjective/qualitative descriptions of male alarm calls in relation to those of females (i.e., the ‘alarm bark’ versus ‘chirp’ distinction, respectively, see **Chapter 3**) suggest that structural differences in the calls of males and females may exceed the associated variation in body size expected if the relationship between the former and the latter was entirely isometric.

In many ways, structural characteristics of the male’s alarm call links them to the “loud calls” of other primate species, which are generally loud, low-frequency calls produced exclusively or preferentially by adult males, and are thought to function in both inter- and intra- group communication (Gautier and Gautier-Hion 1977, Delgado 2006). A number of non-mutually exclusive hypotheses as to the selective pressures underlying the low frequency content of long calls have been proposed. As mentioned above, long calls are often thought to be under strong sexual selection to effectively “exaggerate” a caller’s body size, thus allowing callers to intimidate rivals and/or attract mates (Delgado 2006, Fitch 1999, Charlton et al. 2007a,b, Taylor and Reby 2010). Additionally, the low frequency content of long calls is thought to be resistant to attenuation, allowing calls to travel further without degradation (Mitani and Stuht 1998).

Although the alarm calls of vervet monkeys have been the subject of a great deal of research (Cheney and Seyfarth 1980, Seyfarth et al. 1980a,b, Price et al. 2015), the majority of this work has largely focused on their putative informational content from a referential standpoint, often at the expense of objectively describing, quantifying and investigating the functional significance of other possible sources of structural variation in these calls. Here, I test the hypothesis that sex differences in the alarm calls of adult vervet males and females exceed those expected by proportionate differences in body size. Thus, my null hypothesis is that sex differences in calls are directly proportional to differences in body size. These hypotheses generate testable predictions regarding the relationship between call structure and the body size of callers. For instance, if the male's alarm call has been under specific selection to either exaggerate their body size, or increase the call's effective range relative to the calls of females, I predict that the *relationship* between body size and the acoustic parameters of alarm calls will be different for both sexes. Specifically, I would predict that as males grow, measurements relating to the frequency content of their alarms will be hyperallometric (i.e., the rate of change will be exaggerated) relative to those of females, with males showing a stronger negative relationship between body weight and frequency content than females. Similarly, the duration of alarm calls (and/or the elements that make up the alarm calls) will increase at a higher rate in males relative to females.

Alternatively, if sex differences in call structure are simply a by-product of sexual dimorphism in body size, I predict that the relationship between body size and the acoustic structure of alarm calls will be similar in males and females. Specifically, I predict that the frequency content of alarms produced by males and females will decrease with body size at the same rate. This being the case, while the frequency content of alarm calls should decrease in both males and females as they grow, the larger body size of adult males would result in lower frequency

content in their calls as a simple by-product of larger body size. Similarly, the temporal characteristics of these calls should increase with body size at the same rate in both sexes, with sex differences in adults emerging as a product of body size dimorphism.

Additionally, if the alarm calls of males are under specific selection for longer durations or lower frequency content than would be expected based on body size alone, I expect that sex differences in vocal allometry will be specific to this call type. Differences in body size should theoretically influence the structure of different call types in similar ways. Thus, if sex differences in call structure are simply a by-product of sex differences in body size (i.e., neutral vocal allometry), all call types should be affected in similar ways. If, however, the male's alarm call has been the target of specific selection for lower frequency content or longer duration, I expect that sex differences in the relationship between body size and call structure will be limited to alarm calls.

In order to distinguish between these two hypotheses, I explore the relationship between body weight (as a proxy for body size) and the acoustic structure of the vervet monkey's terrestrial alarm calls (i.e. those given to terrestrial predators) in males and females. Specifically, my goal is to determine whether the relationship between body size and the acoustic structure of terrestrial alarm calls are the same for males and females. Additionally, in order to determine whether the observed sexual dimorphism in call structure is specific to terrestrial alarm calls, I also investigate the relationship between body size and the acoustic structure of another vocalization produced by males and females: the grunt. Grunts are typically low amplitude calls used in close range communication (**Chapter 2**), meaning they should not be under selection for long range transmission. Additionally, that grunts are produced in close range situations suggests that selection for acoustic "exaggeration" of body size through the use of grunts is unlikely, as at close

ranges, receivers could simply validate caller body size through visual channels. As such, I compare the relationship between body size and acoustic structure in grunts and alarm calls, under the assumption that if the male’s terrestrial alarm call has diverged in structure from those of females due to selection for either long range transmission, or body size exaggeration, there should be sex differences in the relationship between body size and the acoustic structure of alarm calls, but not grunts.

4.2 Methods

Vocal data collection in the field

The methods used to record grunts and alarm calls are described in detail in **Chapters 2** and **3** respectively. In short, both grunts and alarm calls were recorded as they occurred naturally in the field. Additionally, because alarm calls are produced relatively rarely under natural conditions, I supplemented naturally occurring calls with those elicited by experimental predator mount presentations (see **Chapter 3**). A breakdown of the number of each call type across all age/sex classes used for this analysis is provided in Table 4.1.

Table 4.1: Number of calls used in the analysis broken down by age/sex class

Call Type	Sex	Infant	Yearling	Juvenile	Sub Adult	Adult
Alarm Call	Male	0	24	96	114	117
	Female	0	5	60	27	60
Grunt	Male	83	51	85	95	211
	Female	10	73	172	157	274

Acoustic analysis

Individual grunts and alarm calls were often produced in bouts of multiple calls. I defined a bout as a cluster of one or more similar calls emitted under one clear context, and separated from other bouts by a time interval longer than any time interval between calls within the bout. For alarm calls, all vocalizations produced in response to a single encounter with a predator or during a single predator mount presentation were considered as belonging to the same bout.

Both terrestrial alarm calls and grunts were often made up of multiple elements (Fig. 4.1). For temporal measurements I included all the elements within a given call. Conversely, variables relating to the frequency content of signals were taken from a single call element. As the first and last elements within grunt vocalizations were often much lower in amplitude than the other elements, I always choose an element from the middle of the call when there were more than two elements. When a call consisted of only two elements, I defaulted to using the first of these. Alarm calls produced by adult males often consisted of a combination of audible inhalatory and exhalatory components (Price et al. 2015, **Chapter 3**). Alarm calls produced by adult females and juveniles could also contain audible inhalation components, though this was less common and, when they did, the inhalation was considerably quieter than the exhalation. For consistency, I conducted spectral analysis only on the exhalatory components of these signals. Unlike grunts, the first exhalatory components of alarm calls were not quieter than any other element and, as such, I used those first elements for spectral analysis.

I selected for acoustic analysis calls that were clear of background noise and had a high signal to noise ratio. Before analysis, calls were high-pass filtered at 100Hz to remove any noise below the lowest frequency of the call (Raven Pro, v1.5). I labeled each call element and saved

them as separate files. Call elements were padded with 0.2 second silent margins using Avisoft SASLab Pro, before being resampled to 16kHz. I performed a Fast Fourier transform (Hamming window function, 1024 points, 93.75% overlap) to generate a spectrogram for each call element. The spectrograms were then exported and analyzed using a custom sound analysis software developed by K. Hammerschmidt (LMA - Lautmusteranalyse v. 2018_0.4).

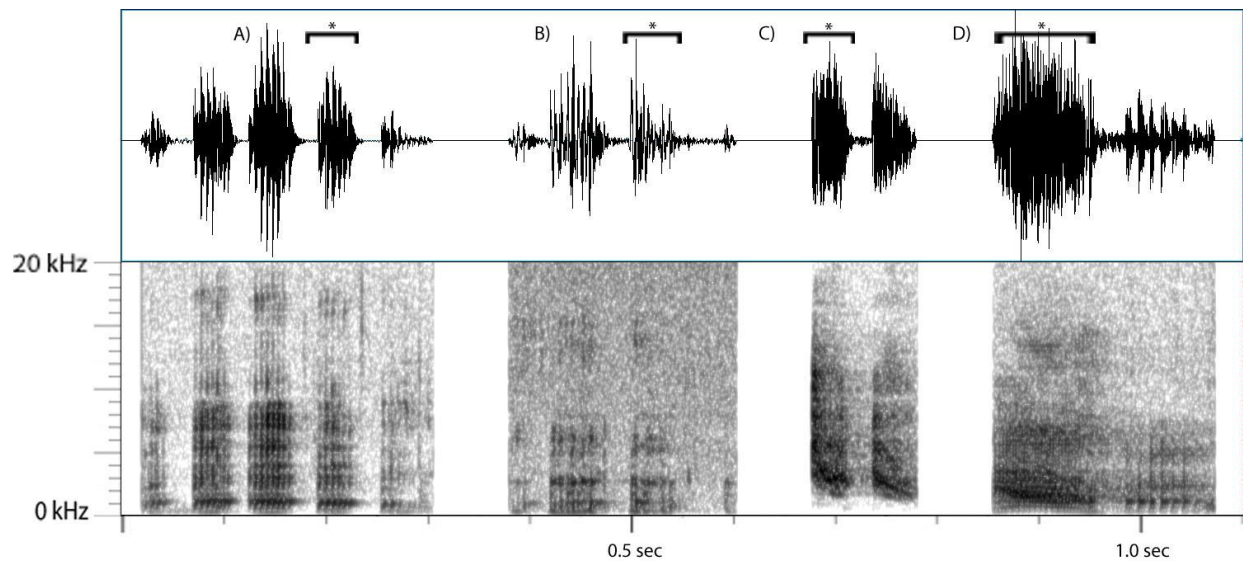


Figure 4.1: Oscillograms and spectrograms representing A) a grunt produced by an adult female, B) a grunt produced by an adult male, C) a terrestrial alarm call (chirp) from an adult female, and D) A terrestrial alarm call (Threat alarm bark) from an adult male. In each call, one element is highlighted with an Asterisk (*).

Variable selection

For this analysis, I selected acoustic variables that a) were rated as being relatively important for discriminating between calls produced by either males or females for either grunts or alarm calls (see **Chapter 2** and **3**) and, b) were likely to vary with caller body size based on the source filter framework of vocal sound production (Table 4.2). For temporal measurements, I chose call duration and mean element length, which were both rated as being important for

discriminating between the calls of males and females for grunts. From a theoretical perspective, both these variables are likely to vary with a caller's lung size, with larger lungs being able to produce longer calls, and longer call elements. To represent the frequency content of the signals, I chose the mean first dominant frequency band (DF1), and the second distribution of the frequency amplitudes (DFA2). Although DF1 did not differ significantly between males and females in grunt vocalizations (see **Chapter 2**), it was very important for distinguishing between the alarm calls of males and all other age-sex classes. DFA2 was rated as being the most important variable for distinguishing between the grunts of males and females, and was also rated as being highly important for distinguishing between male and female alarms.

Table 4.2: Acoustic Variables used in the analysis

Temporal measurements – Based on entire call	
FullCallDur	Duration of call [ms]
Mean element length	Mean element length within a call [ms]
Spectral measurements – Based on single call element	
DFA2	Distribution of the frequency amplitudes 2: Mean second quartile of amplitude in spectrum [Hz]
DF1	Mean frequency of 1 st dominant frequency [Hz]

Body mass

Mass data were collected in one of two ways. First, a field scale (Adam Equipment CPW Plus-15 weighing scale) was brought into the field on a regular basis to take weights of individuals opportunistically. The scale was fitted with a wooden platform (70x70cm) to increase the area on which the monkeys could stand. The wooden platform was baited with a small amount of honey,

and/or by attaching small objects, including mirrors or baby toys, that were designed to be interesting to the monkeys. Monkeys could freely approach or leave the scale of their own volition. Using the weight data collected from the scales, the weight of males and females could be modeled for each day (Jarrett 2019), providing us with the ability to assign a predicted weight for callers based on their ages (in days) when vocalizations were recorded. However, as many of my samples were recorded from individuals for which I had inadequate weight data with which to model weight, I modeled the weight of males and females respectively for this analysis, and not the weight of specific individuals (Fig. 4.2A).

As adults were generally uninterested in the portable field scale, I did not model weight data for adult callers. A number of adults at the field site were, however, implanted with temperature loggers, which are used to obtain body temperature measurements (Lubbe et al. 2014, McFarland et al. 2015). The annual removal or replacement of loggers required that the animals be tranquilised and, as part of this procedure, biometric data, including body weight, were obtained from each animal. For my analyses, I averaged the body weights taken from each individual since the surgeries began in 2010 (Fig. 4.2B).

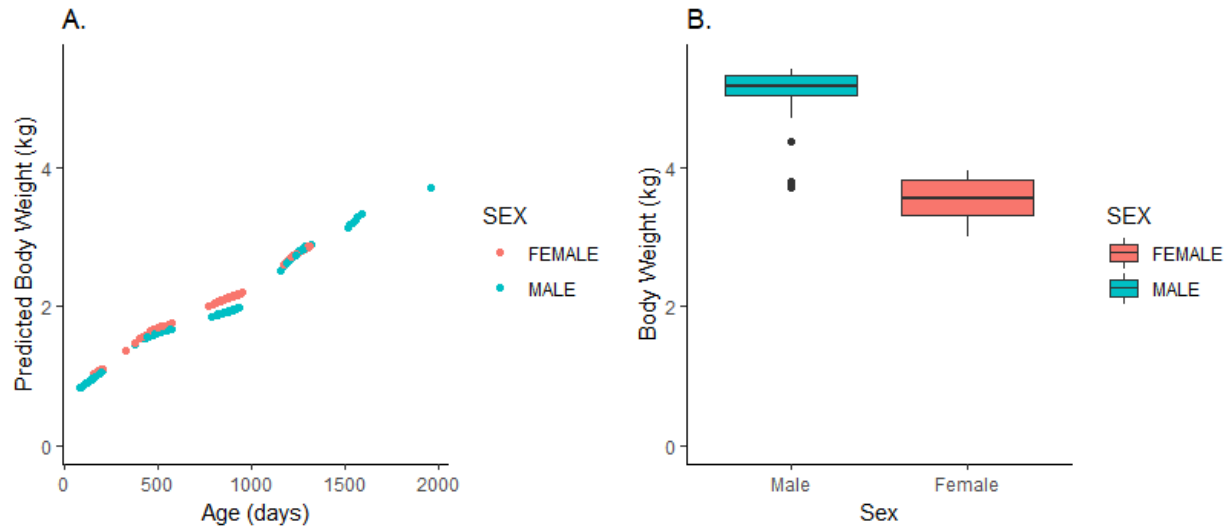


Figure 4.2: Weight data used to model the interaction between body weight and each acoustic parameter. Scatterplot A) represents modeled weight data for juvenile males and females. Each point represents weights predicted for either a male (blue) or female (red) juvenile caller on the day that its call was recorded. Models were based on data collected in the field using a portable scale. Boxplot B) represents the average weight for adult males and females based on biometric data collected during annual surgeries since 2010.

4.3 Data Analysis

Models

I modeled the effects of call type (categorized by sex), body weight, and their interaction on each of the frequency related acoustic variables (DF1, DFA2) using two linear mixed effects models. I modeled the relationship between both temporal variables (Mean element length, and call duration) and body weight using two Generalized linear mixed models, using a log normal distribution. All models were run within a Bayesian framework using the ‘brms’ package (Bürkner 2017). I choose a Bayesian framework as I wanted to model both how the mean and the variance for each acoustic parameter changed with body weight. For each model, I specified four chains and 1000 iterations. All models used weakly informative priors centered on zero. For all models,

\hat{R} values were equal to or below 1.02, indicating model convergence, and the effective sample size was always above 325.

For each response variable (DF1, DFA2, Call duration, Element Length), I used body weight, call type (Female Grunts, Male Grunts, Female Alarm, Male Alarm), and their interaction, as fixed effects for both the mean and the variance. For both models related to the frequency content of the calls (DF1, DFA2), I standardized and centered the response variable. For the models related to temporal parameters (Call duration, Element Length), I standardized the response variables, but did not center them, as the log normal distribution does not allow for negative values. The estimation of body weight was centered for all analyses. To control for the models containing more than one call per individual, and more than one call per calling bout, caller ID and calling bout were entered as crossed random effects and specified as random intercepts in all four models.

When interpreting my results, I considered the uncertainty in sign, and magnitude, for each model parameter (Gelman and Carlin 2014). If the 95% credible intervals (CIs) around the estimate were positive, I interpreted that as an indication that the effect was positive (and negative if the CIs were negative). If the range of the CIs included zero, I interpreted this as an indication that the sign could not be clearly distinguished. I used the range of the CIs as an indication of the certainty of the magnitude of the effect; wider CIs suggest a higher degree of uncertainty surrounding the magnitude. I used the models to predict the interaction between body weight and both the mean, and variance of each acoustic parameter. Using this method, I was able to compare the rate of change for each parameter in relation to increasing body weight.

To make posterior predictions (Full model predictions) on the scale of the response variables, I used my models to make predictions about how the means, and variance around these means, changed with body weight. To determine whether the acoustic parameters of males and

females were affected differently due to weight within call types, I calculated the differences in the posterior predictions. This allowed me to determine whether the acoustic structure of calls produced by males and females responded differently to increased body weight.

As variables relating to the frequency content of the signal are expected to vary based on the length of the vocal folds and the length of the supra-laryngeal vocal tract, I used the log₁₀ of body weight (simply body weight from here on) in the models, to account for the cubic relationship between volume and a linear dimension.

4.4 Results

Spectral parameters

The results of my linear mixed models for each of the spectral acoustic variables are provided graphically in Fig. 4.3A/B and Fig. 4.4A/B (as well as in Table A.2 and Table A.3 in the appendix to chapter 4). The R^2 values for both models were relatively high, with the absence of much influence of the random effects on the overall outcome, thus indicating that the main effects were able to capture most of the variance effectively (DF1 - R^2 conditional = 0.76, SE: 0.00, R^2 marginal = 0.76, SE: 0.00; DFA2 - R^2 conditional = 0.75, SE: 0.01, R^2 marginal = 0.68, SE: 0.01).

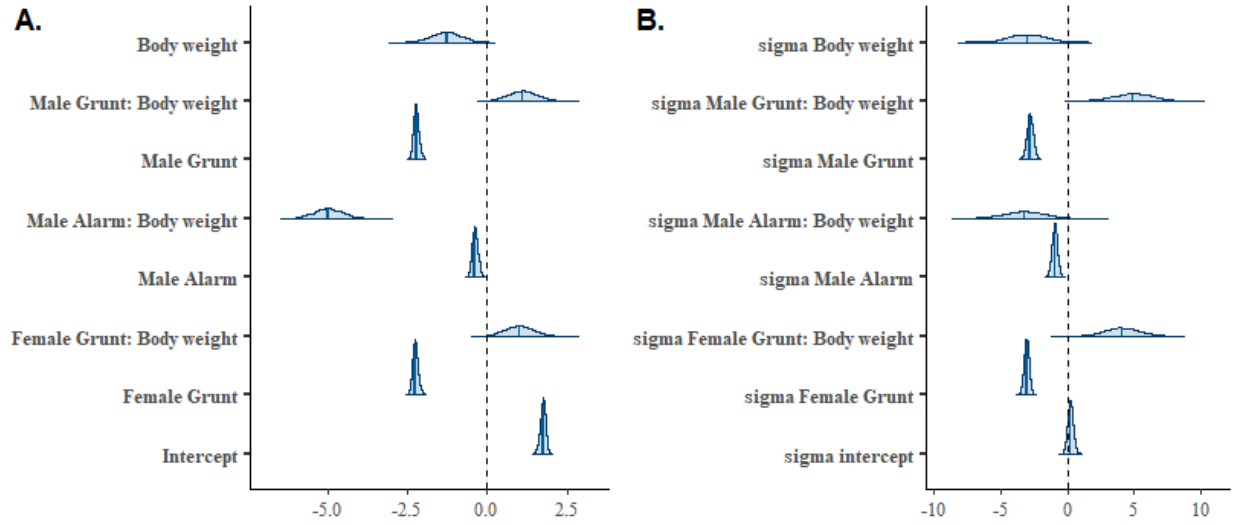


Figure 4.3: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and DF1 in grunts and alarms produced by males and females, and (B) the relationship between body weight and variability in DF1 in grunts and alarms produced by males and females. Estimates for call types are relative to female alarm calls.

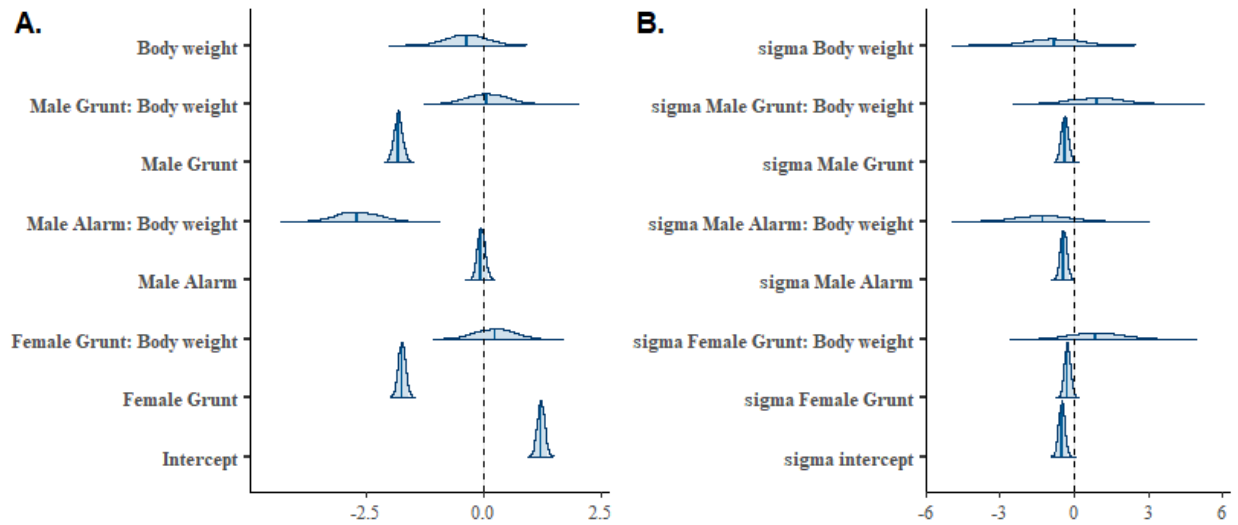


Figure 4.4: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and DFA2 in grunts and alarms produced by males and females, and (B) the relationship between body weight and variability in DFA2 in grunts and alarms produced by males and females. Estimates for call types are relative to female alarm calls.

DF1: Alarm Calls

I found a relatively strong negative relationship between body mass and the mean first dominant frequency band (DF1) for alarm calls produced by males (Estimate = -500.54Hz/kg, CI: -519.81, -481.30Hz/kg, Fig. 4.5/Fig. 4.6a). The alarm calls produced by females showed a similar trend, in that larger subjects had lower first dominant frequency bands, though the credible intervals surrounding the estimate were relatively wide, suggesting a higher range of uncertainty surrounding the precise magnitude of this effect (Estimate = -102.78 hz/kg, CI: -170.54, -38.34Hz/kg Fig. 4.5). The model revealed clear sex differences in terms of the relationship between body weight and DF1 for alarm calls; the inverse relationship between body weight and DF1 was clearly stronger in males than it was for females (Df1 males- Df1 Females: Estimate = -397.76Hz/kg, CI: -470.40, -333.86 Hz/kg, Fig. 4.2a, Fig. 4.5/Fig. 4.6a). The model also showed that larger individuals (both males and females) had less variability in DF1 (Males: -0.82Hz/kg, CI: -0.99, -0.66; Females: -0.39Hz/kg, CI: -0.74, -0.01, Fig. 4.7), though this effect was again stronger for males than it was for females (σ Df1 males – σ Df1 Females: Estimate = -0.42, CI: -0.87, -0.05, Fig. 4.7).

DF1: Grunts

In both males and females, larger subjects also produced grunts with lower DF1 values (Males: Estimate = -15.02Hz/kg, CI: -20.70, -9.15 Hz/kg - Females: Estimate = -22.90 Hz/kg, CI: -32.70, -12.67, Fig. 4.5/4.6b). Unlike alarm calls however, there was no evidence of clear sex differences in the relationship between body weight and DF1 (Df1 males- Df1 Females: Estimate = 7.88 Hz/kg, CI: -4.67, 18.86, Fig. 4.5). Similarly, there was no evidence of sex differences in the

relationship between body size and variation in this parameter (Σ Df1 males - Σ Df1 Females: 0.134 Hz/kg, CI: -0.21, 0.43, Fig: 4.7).

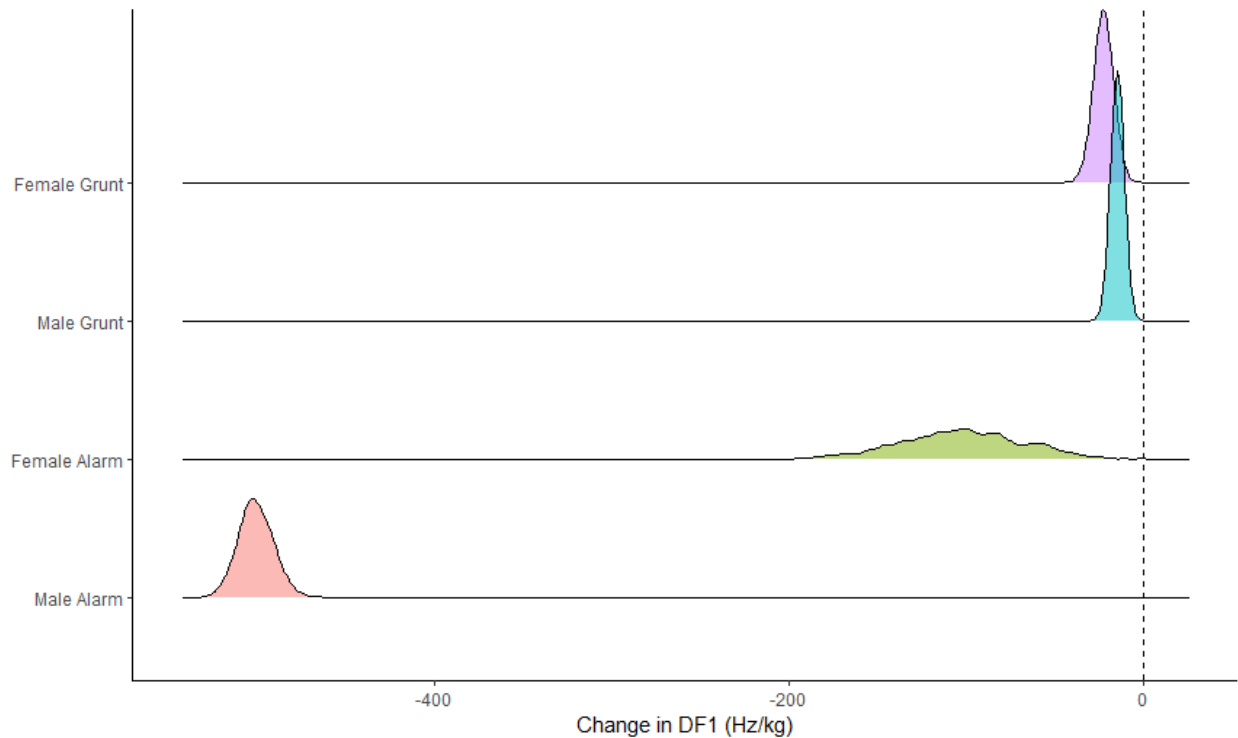


Figure 4.5: Density plots represent the rate of change in DF1 as body size increased (Hz/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). Values below zero (dotted line) suggest that as body size increased, DF1 decreased. The magnitude of this effect is represented by the x-axis. Wider distributions suggest the model predictions were less confident in terms of the magnitude. Values are back transformed into the original scale.

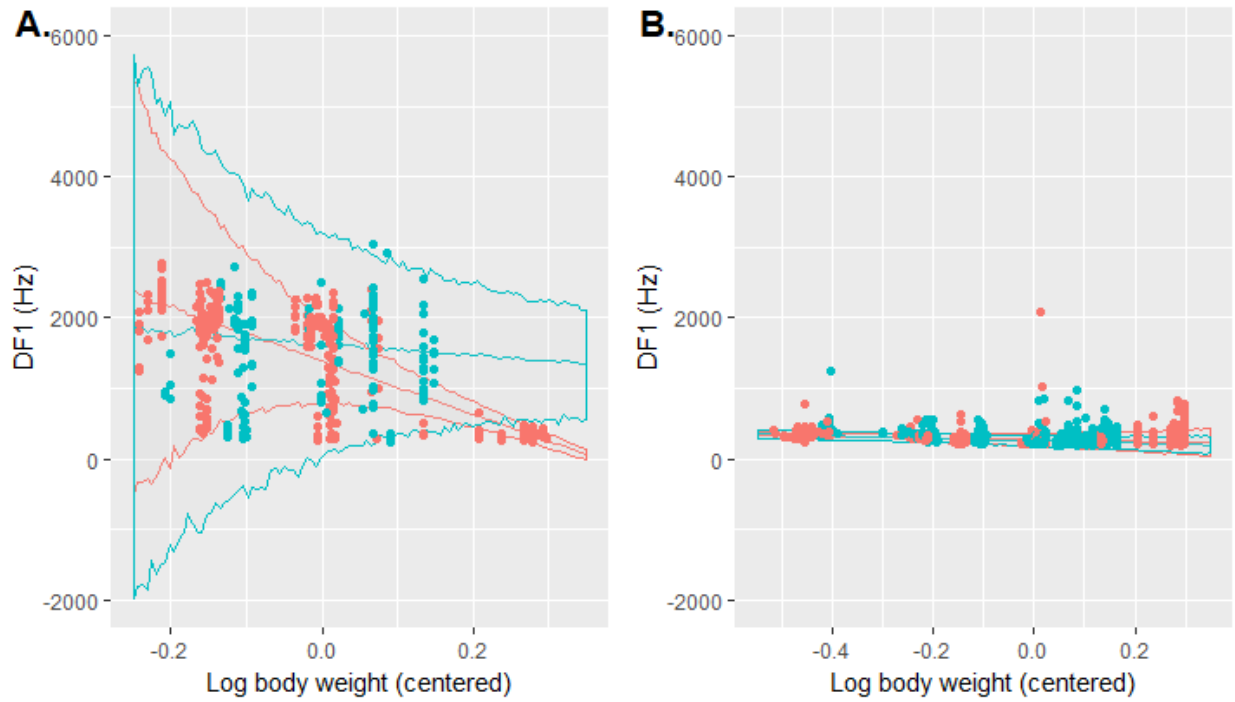


Figure 4.6: Scatterplots illustrating the relationship between body mass and DF1 for alarm calls and grunts. Scatterplot (A) represents the relationship between bodyweight and the lowest dominant frequency band (DF1) for alarm calls. Scatterplot (B) represents the relationship between bodyweight and DF1 for grunts. In each scatterplot, the red line represents the estimated effect for males, while the blue line represents the estimated effect for females.

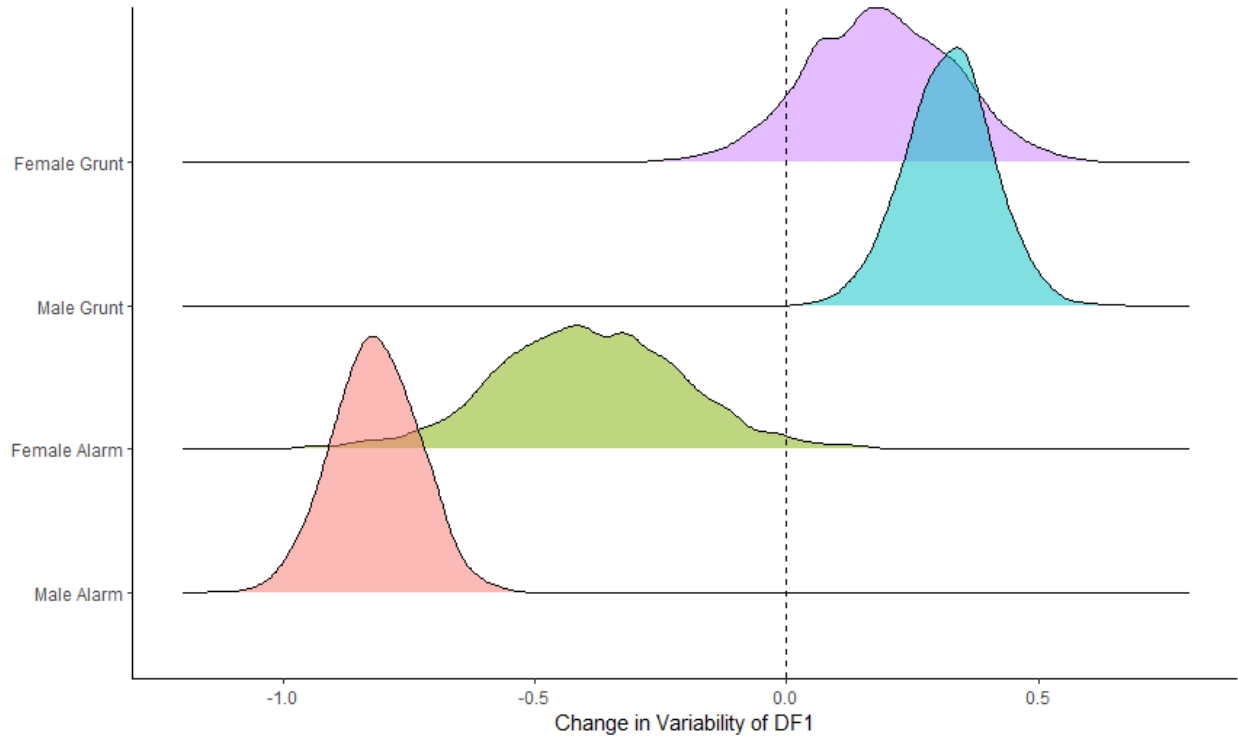


Figure 4.7: Density plots represent the rate of change in the variability of DF1 as body size increased (Hz/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). The magnitude of the effect is represented by the x-axis. Values below zero (dotted line) suggest that as body size increased, variability in DF1 decreased. Values above the dotted line suggest that as body size increased, variability in DF1 increased. Wider distributions suggest the model predictions were less confident in terms of the magnitude. Values are back transformed into the original scale.

DFA2: Alarm Calls

The relationship between body weight and the frequency at which the second quartile of global energy was reached (DFA2) was similar to that between body weight and DF1: For alarm calls produced by males, there was a substantial drop in DFA2 as body size increased (Estimate = -783.59Hz/kg, CI: -813.77, -753.48, Fig. 4.8/4.9a). There was a similar, but weaker negative relationship between this parameter and body weight for alarm calls produced by adult females (Estimate = -160.90 Hz/kg, CI: -266.98, -60.02, Fig. 4.8). The wide confidence intervals

surrounding this estimate suggest a higher range of uncertainty surrounding the precise magnitude of this effect relative to males (Fig. 4.8). The model revealed sex differences in the relationship between body weight and DFA2 for alarm calls; DFA2 dropped for males disproportionately with increasing body size relative to females (DFA2 males - DFA2 Females: Estimate = -622.43 Hz/kg, CI: -736.43, -522.67, Fig. 4.2b, Fig. 4.8/4.9a). I did not find evidence of sex differences in the relationship between body size and variability in DFA2 for alarm calls (Estimate = -0.17, CI: -0.42, 0.12, Fig. 4.10).

DFA2: Grunts

There was a negative relationship between DFA2 and body size in grunts for both males (Estimate = -30.21 Hz/kg, CI: -41.63, -18.41, Fig. 4.8/4.9b) and females (Estimate = -46.07 Hz/kg, CI: -65.79, -25.50). As was the case for DF1, there was no evidence of sex differences in the interaction between body size and DFA2 for grunts (DFA2 males- DFA2 Females: Estimate = 15.86 Hz/kg, CI: -8.99, 37.34, Fig. 4.8). I did not find evidence of sex differences in the relationship between body size and variability in DFA2 for grunts (Estimate = -0.00020 Hz/kg, CI: -0.19, 0.20, Fig. 4.10).

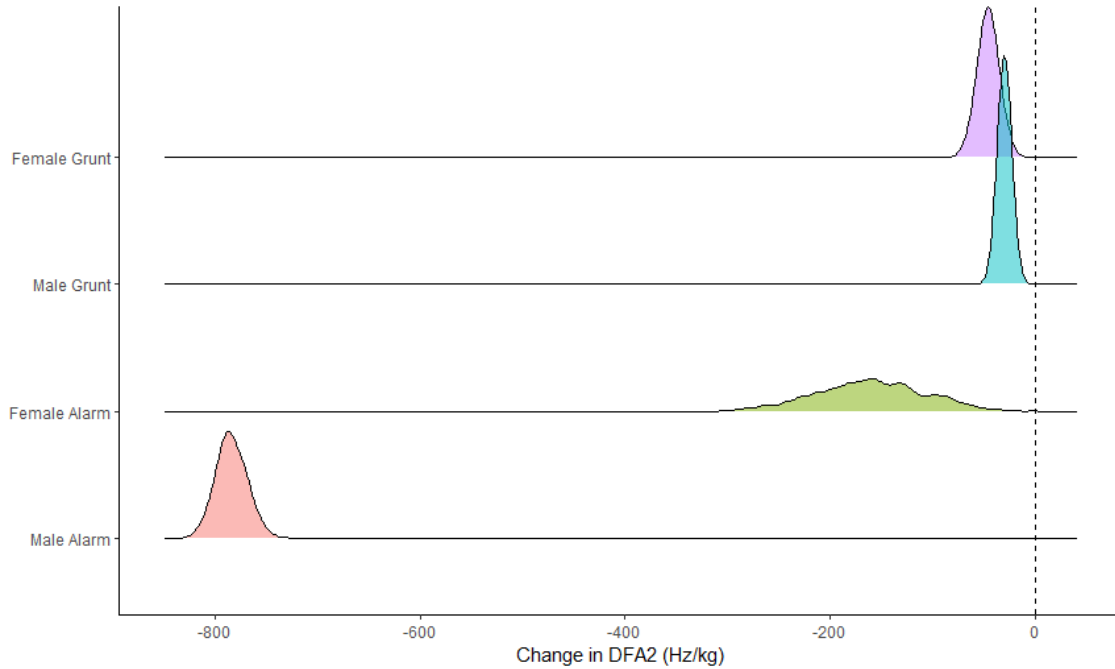


Figure 4.8: Density plots represent the rate of change in DFA2 as body size increased (Hz/kg) predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). Values below zero (dotted line) suggest that as body size increased, DFA2 decreased. The magnitude of this effect is represented by the x-axis. Wider distributions suggest the model predictions were less confident in terms of the magnitude. Values are back transformed into the original scale.

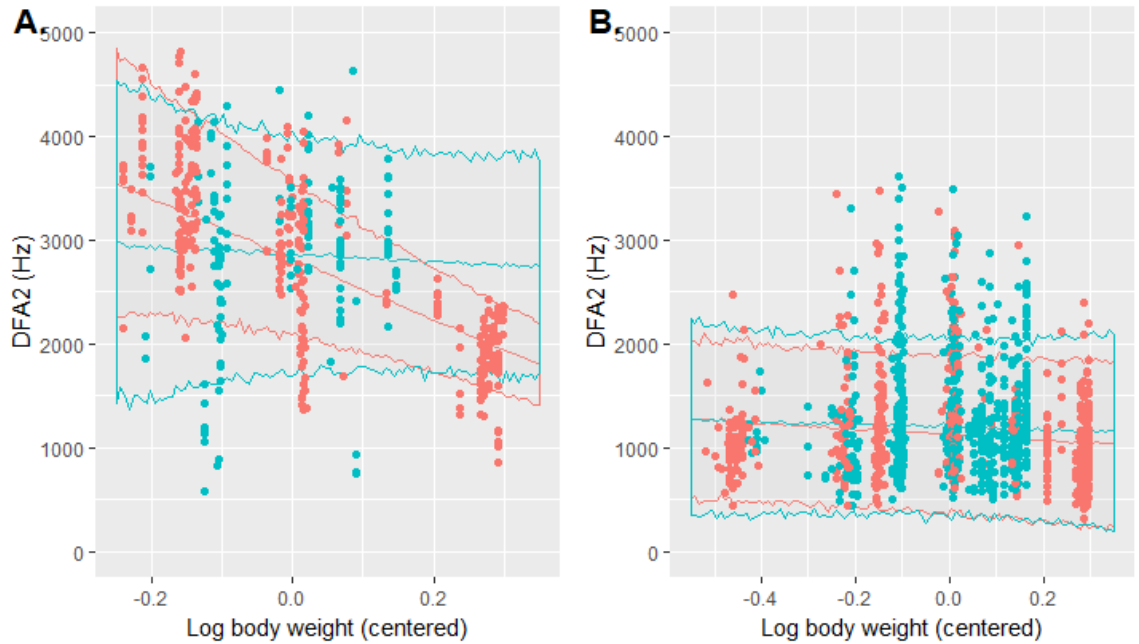


Figure 4.9: Scatterplots illustrating the relationship between body weight and DFA2 for alarm calls and grunts. Scatterplot A) represents the relationship between bodyweight the frequency at which the second quartile of global energy was reached (DFA2) for alarm calls. Scatterplot B) represents the relationship between bodyweight and DFA2 for grunts. In each scatterplot, the red line represents the estimated effect for males, while the blue line represents the estimated effect for females.

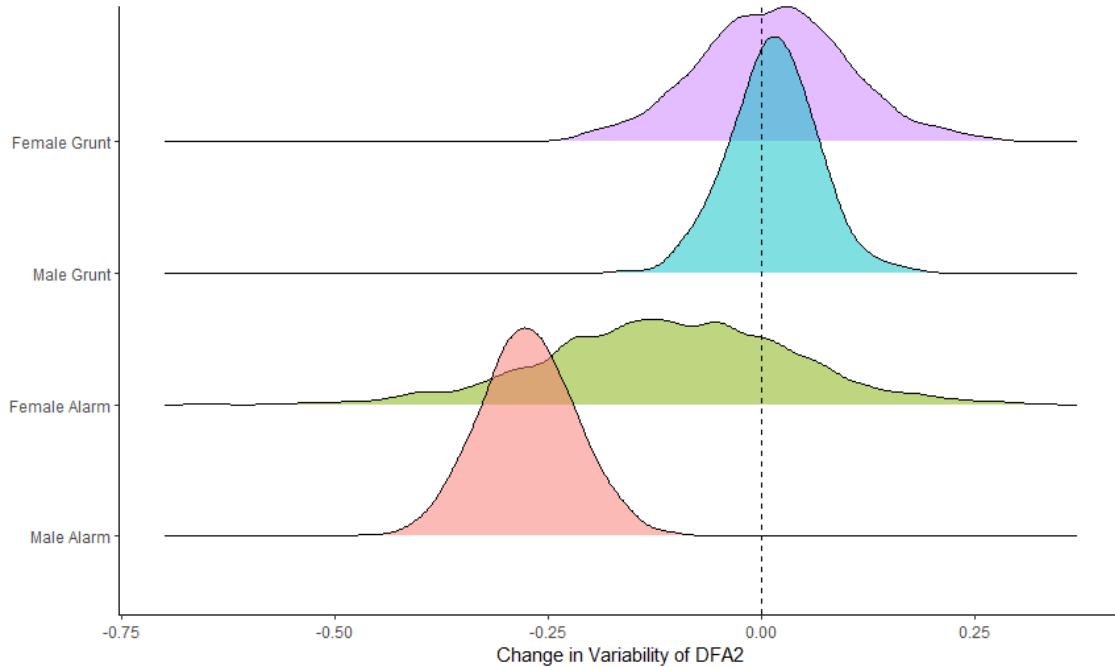


Figure 4.10: Density plots represent the rate of change in the variability of DFA2 as body size increased (Hz/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). The magnitude of the effect is represented by the x-axis. Values below zero (dotted line) suggest that as body size increased, variability in DFA2 decreased. Values above the dotted line suggest that as body size increased, variability in DFA2 increased. Overlap with the dotted line suggests that the direction of the effect cannot be distinguished from zero (no effect). Wider distributions suggest the model predictions were less confident in terms of the magnitude. Values are back transformed into the original scale.

Temporal parameters

The results of my linear mixed models for each of the temporal acoustic variables are provided in Fig. 4.11A/B and 4.12 A/B (as well as in Table A.4 and Table A.5 in the Appendix). My R^2 values for both models dealing with temporal parameters were relatively low, highlighting the fact that both models were limited in the amount of variance in the dependant variable they could explain (Call duration - R^2 conditional = 0.24, st. error: 0.02; R^2 Marginal = 0.17, st. error: 0.01; Element length - R^2 conditional = 0.35, st. error: 0.05; R^2 Marginal = 0.21, st. error: 0.02).

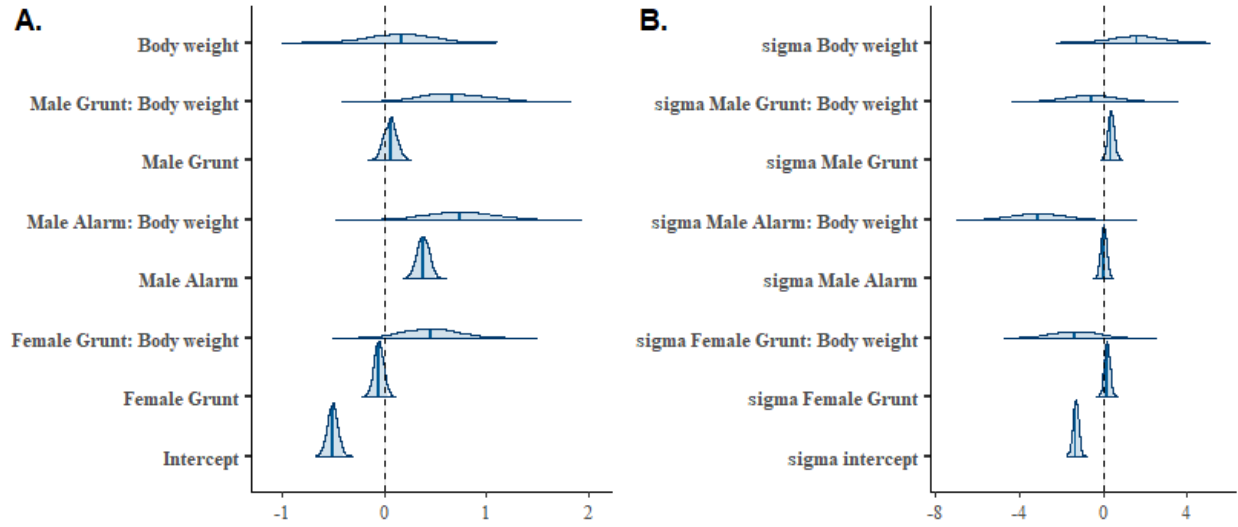


Figure 4.11: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and mean element length in grunts and alarms produced by males and females, and (B) the relationship between body weight and variability in the mean element length in grunts and alarms produced by males and females.

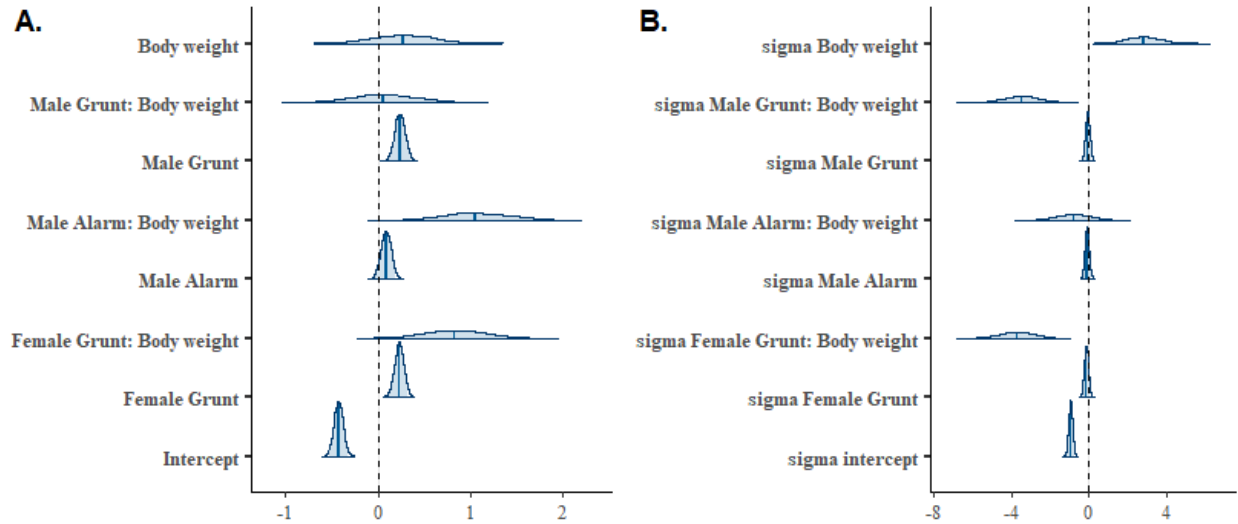


Figure 4.12: Density plots representing estimates (standardized beta coefficients) of (A) the relationship between body weight and call duration in grunts and alarms produced by males and females, and (B) the relationship between body weight and variability in call duration in grunts and alarms produced by males and females.

Element length: Alarm calls

My model suggests that there was a positive relationship between body size and element length for alarm calls produced by males (Estimate = 7.60 ms/kg, CI: 3.65, 11.36, Fig. 4.13/4.14a). Although my model level estimates do not give clear evidence of an association between body size and element length for females' alarm calls (Estimate = 2.48 ms/kg, CI: -2.48, 7.54), there seemed to be some indication of a weak positive association (Fig. 4.13). Although I did not find clear evidence for sex differences in the relationship between body size and element length for alarm calls (male – female: 5.13 ms/kg, CI: -0.92, 10.94, Fig. 4.13), I did find sex differences in the relationship between body size and the amount of variation for this parameter; while variability in the length of the elements that made up the alarm calls of both males and females decreased with body size (male, Estimate: -6.22, CI: -7.55, -5.06; female, Estimate: -2.99, CI: -5.63, -0.06, Fig. 4.15), this effect was greater for males overall (male – female, Estimate: -3.23, CI: -6.58, -0.37).

Element length: Grunts

With increasing body size, both males and females produced grunts that were composed of longer call elements (Males: Estimate = 7.61ms/kg, CI: 4.87, 10.70; Females: Estimate = 4.09ms/kg, CI: 1.18, 7.59, Fig. 4.13/4.14b). I did not find evidence for sex differences in the relationship between body size and element length in grunts (element length males – element length females: -1.62ms/kg, CI: -6.05, 3.77, Fig. 4.13). Similarly, I found no evidence for sex differences in the effect of body size on variability of element length (Estimate: -0.80, CI: -2.58, 1.24, Fig. 4.15).

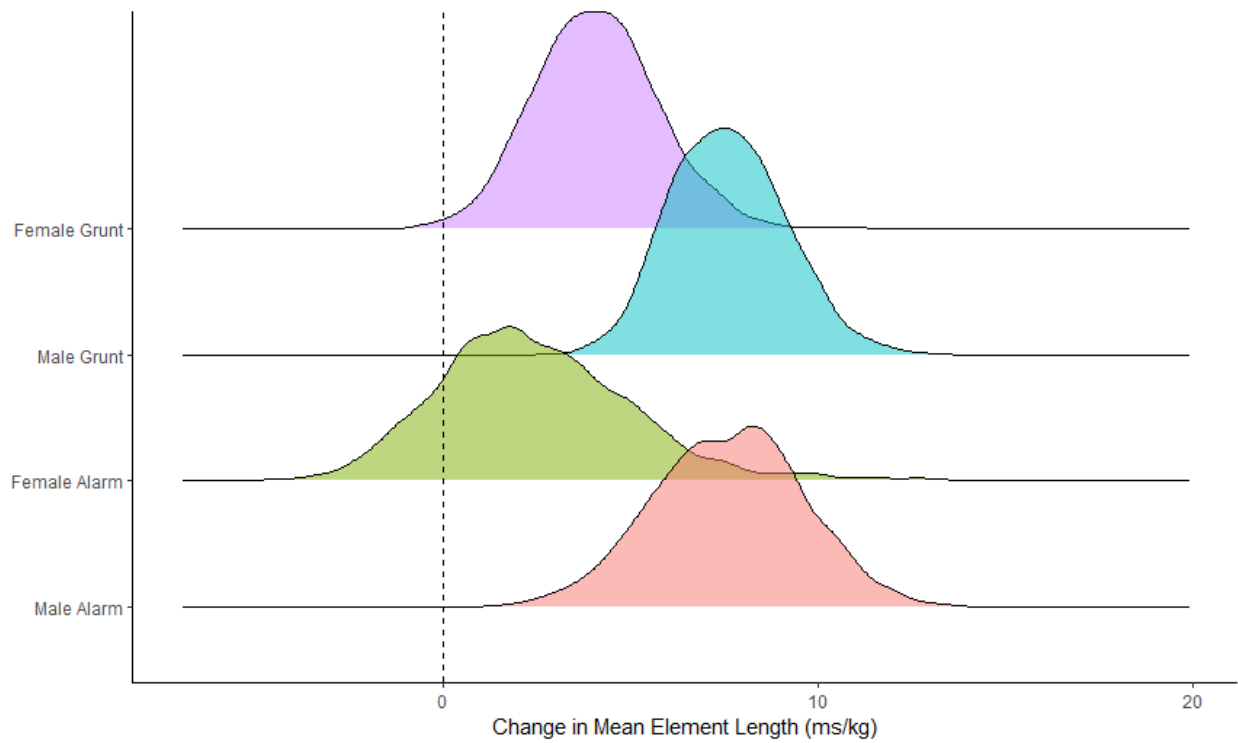


Figure 4.13: Density plots represent the relationship between mean element length and body size (ms/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). Values above zero (dotted line) suggest that as body size increased, so too did mean element length. Overlap with the dotted line suggests uncertainty in the direction of the effect. The magnitude of the effect is represented by the x-axis. Wider distributions suggest the model predictions were less confident in terms of the magnitude of the effect. Values are back transformed into the original scale.

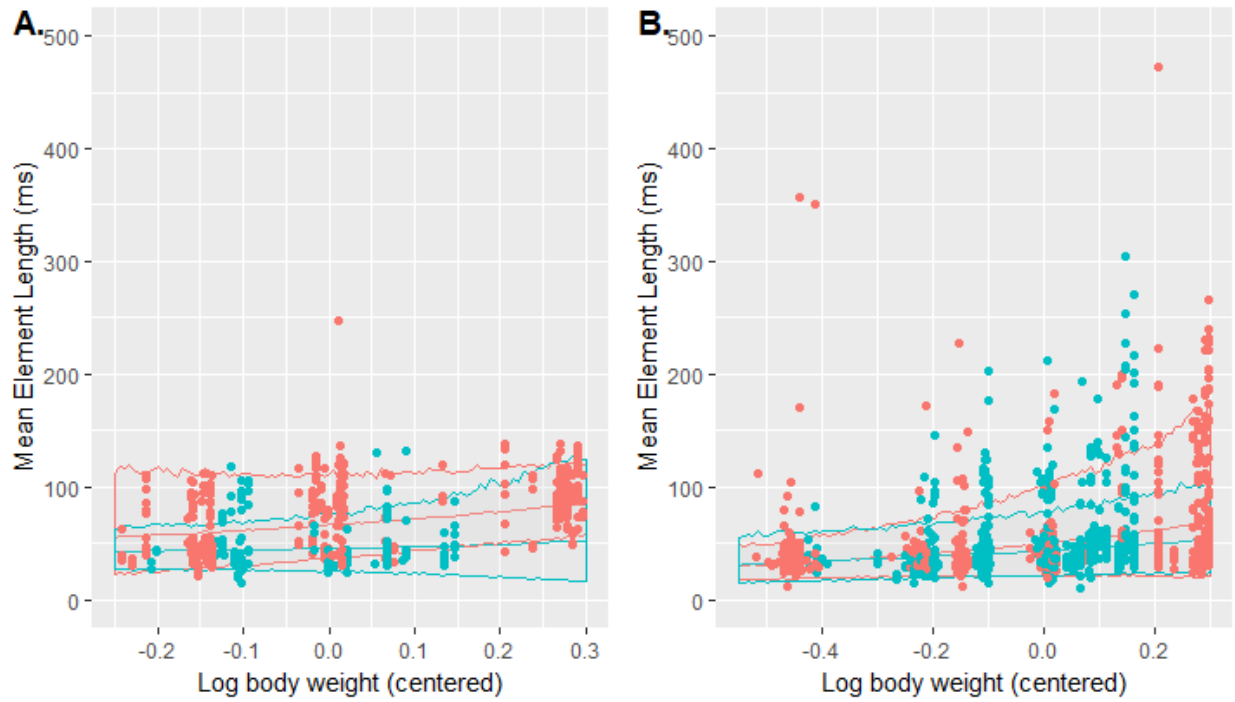


Figure 4.14: Scatterplots illustrating the relationship between body weight and mean element length for alarm calls and grunts. Scatterplot A) represents the relationship between bodyweight the mean element length for alarm calls. Scatterplot B) represents the relationship between bodyweight and mean element length for grunts. In each scatterplot, the red line represents the estimated effect for males, while the blue line represents the estimated effect for females.

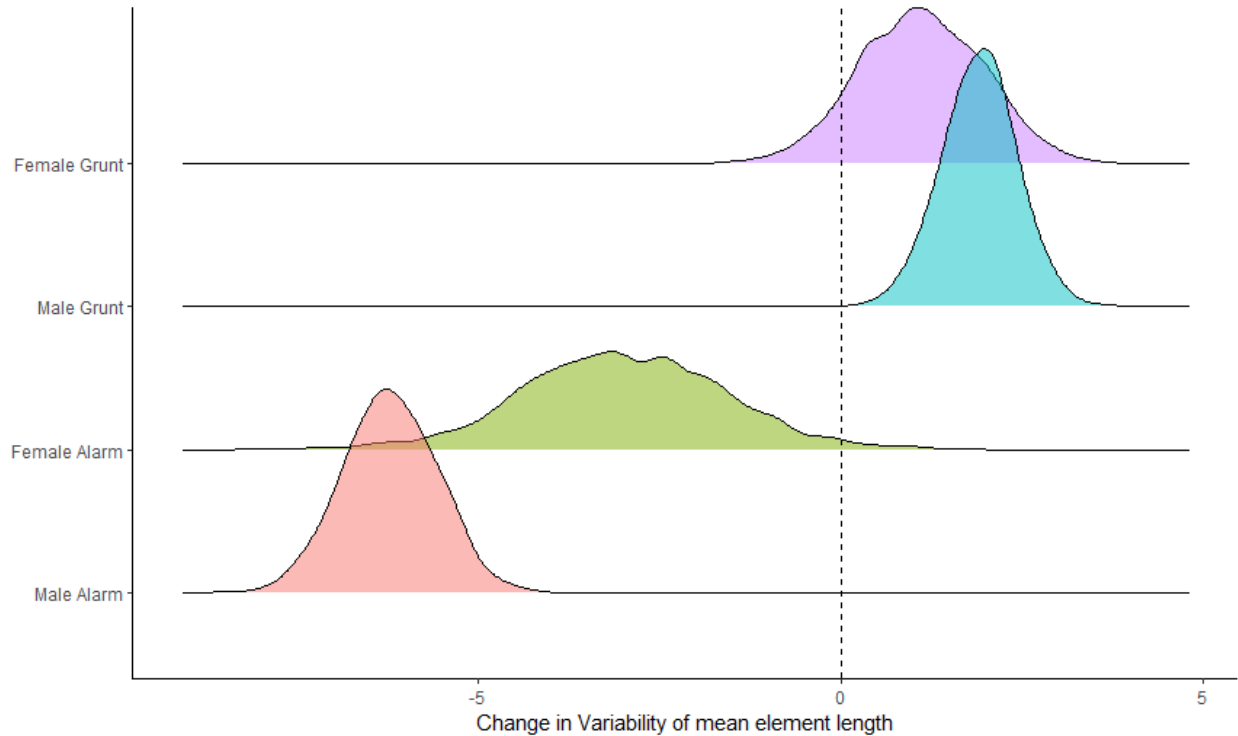


Figure 4.15: Density plots representing the relationship between variability in element duration and body size, as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). The magnitude of the effect is represented by the x-axis. Values below zero (dotted line) suggest that as body size increased, variability in element length decreased. Values above the dotted line suggest that as body size increased, variability in element length increased. Overlap with the dotted line suggest uncertainty in the direction of the effect (i.e., the effect can not be clearly distinguished from zero). Wider distributions suggest the model predictions were less confident in terms of the magnitude of the effect.

Call Duration: Alarm calls

My model showed there was a positive association between body size and call length for alarm calls produced by males (Estimate = 42.27 ms/kg, CI: 29.92, 54.65, Fig. 4.16, 4.17a). Conversely, the model did not show clear evidence of a relationship between body weight and call duration for alarm calls produced by females (Estimate = 32.46 ms/kg, CI: -0.89, 79.55, Fig. 4.16). Still, as was the case for element length, the model results do seem to suggest the possibility of a weak interaction between increasing body size and call duration for female alarm calls; although

the range of the estimates did include zero, the range was clearly biased towards positive values. My model suggested that there were no sex differences in the relationship between body size and call duration for alarm calls (male call duration – female call duration: 9.81ms/kg, CI: -36.93, 48.14, Fig. 4.16). I also found no evidence for sex differences in the relationship between caller body size and variability in call length (male-female, Estimate: -0.81, CI: -2.68, 0.90, Fig. 4.18).

Call Duration: Grunts

The model suggests that larger individuals produced grunts that were longer in duration (Males: Estimate = 6.57ms/kg, CI: 0.16, 12.24; Females: Estimate = 26.50ms/kg, CI: 16.42, 38.35, Fig. 4.16/4.17b). My results indicated that the relationship between call length and body weight differed for males versus females; grunts became disproportionately long in females as they grew compared to the grunts of males (Call duration females-call duration males: 19.94ms/kg, CI: 6.89, 32.08, Fig. 4.16). I found no evidence for sex differences in the effect of body size on variability in call duration in grunts (Estimate: -0.25, CI: -1.05, 0.63, Fig. 4.18).

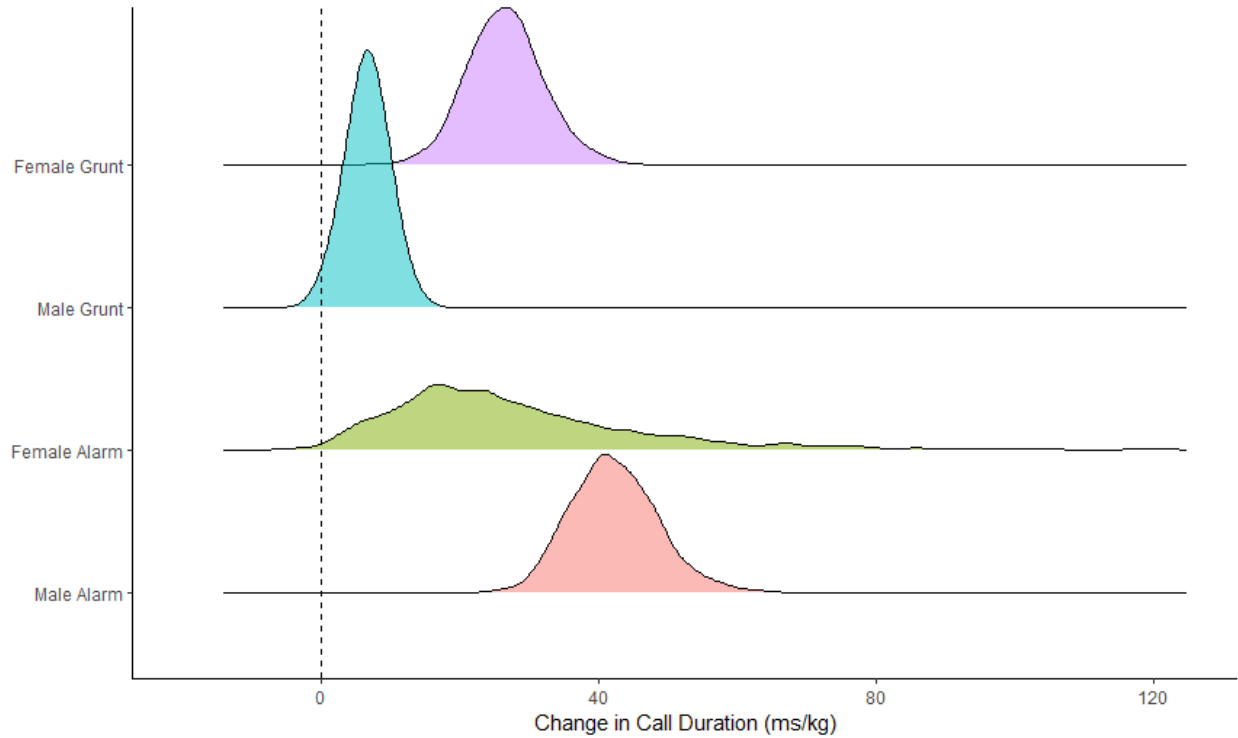


Figure 4.16: Density plots representing the relationship between call duration and body size (ms/kg), as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). Values above zero (dotted line) suggest that as body size increased, so too did call duration. Overlap with the dotted line suggests uncertainty in the direction of the effect. The magnitude of the effect is represented by the x-axis. Wider distributions suggest the model predictions were less confident in terms of the magnitude of the effect. Values are back transformed into the original scale.

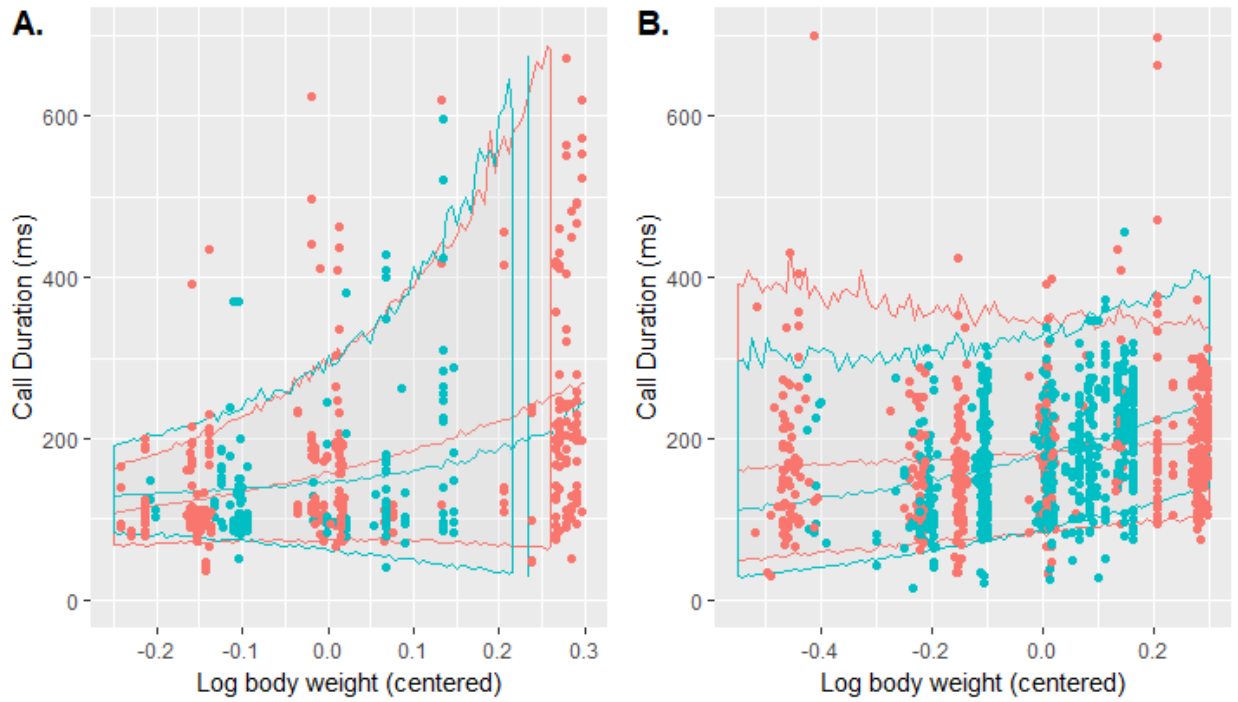


Figure 4.17: Scatterplots illustrating the relationship between body weight and call duration for alarm calls and grunts. Scatterplot A) represents the relationship between bodyweight and call duration for alarm calls. Scatterplot B) represents the relationship between bodyweight and call duration for grunts. In each scatterplot, the red line represents the estimated effect for males, while the blue line represents the estimated effect for females.

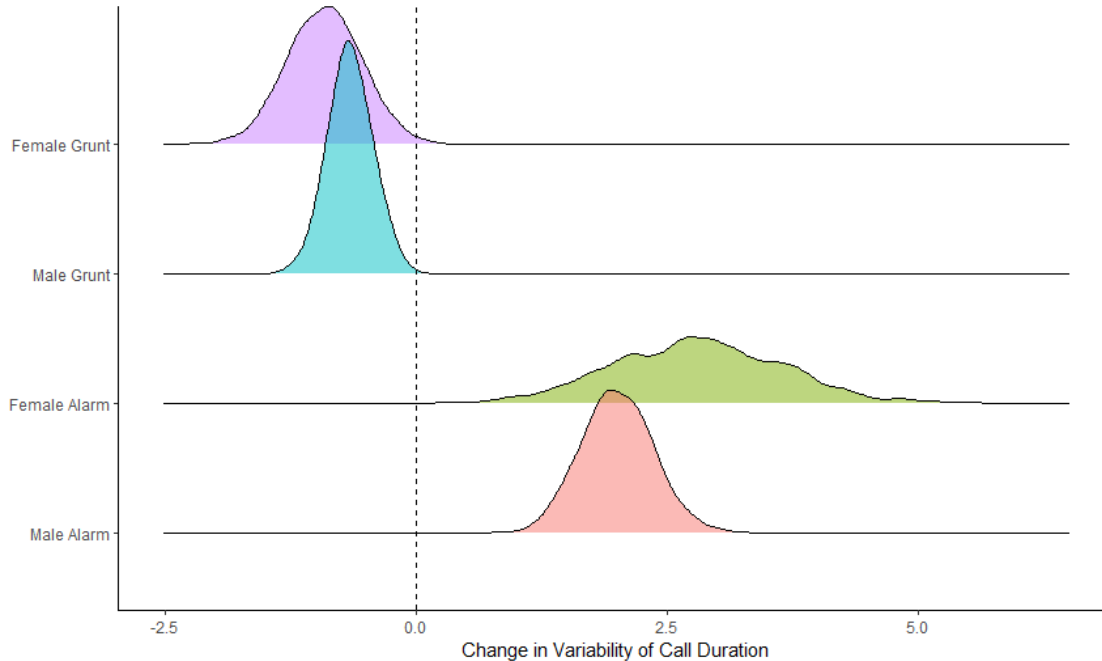


Figure 4.18: Density plots representing the relationship between variability in call duration and body size, as predicted by the model for all call types (Female grunts, Male grunts, Male alarms, Female alarms). The magnitude of the effect is represented by the x-axis. Values below zero (dotted line) suggest that as body size increased, variability in call duration decreased. Values above zero (the dotted line) suggest that as body size increased, variability in call duration increased. Overlap with the dotted line suggest uncertainty in the direction of the effect. Wider distributions suggest the model predictions were less confident in terms of the magnitude of the effect.

4.5 Discussion

My results indicate that both temporal and spectral parameters of grunts and alarm calls varied with caller body size. Specifically, larger individuals tended to produce calls that were longer, were composed of longer elements, had lower first dominant frequencies, and lower frequency distributions compared to smaller individuals. These trends are consistent with what would be expected based on what is known about the mammalian vocal production system. Specifically, as lung and body size are correlated, smaller animals are unable to produce vocalizations with durations that match those of larger individuals (Fitch and Hauser 2003). Additionally, because larger bodied animals have longer vocal folds and longer vocal tracts, the vocal signals they

produce are characterized by energy concentrations at lower frequencies. Because I was not able to measure the fundamental frequency, or position of specific resonances (formants) in my sample, my analysis does not allow me to determine whether the observed spectral differences in these calls are related to changes in the source waveform or the filter function of the vocal tract. Future studies should attempt to distinguish between the influence of the source versus the filter on the vocalizations produced by vervet monkeys. Regardless, the general trends observed here are in line with the interpretation that vocal tract structure plays an important role in the observed body size-related differences in call structure.

I did not find any evidence for sex differences in the relationship between body weight and either of the temporal measurements (mean element length, or call duration) in alarm calls. This suggests that sex differences in the duration of calls or in the durations of the elements that make up the calls at adulthood are likely to be a by-product of sex differences in body size; although calls (and call elements) may be longer in males at adulthood, this is likely due to the fact that males continue to grow after female growth has plateaued (Turner et al. 1997). This dimorphism in body size likely coincides with lung size, allowing males to produce longer calls relative to females of the same age (Fitch and Hauser 2003). I did, however, find that in males, variability in element duration decreased as individuals grew larger, suggesting that element length tends to become more stereotyped in males as they mature

My analysis revealed that for grunt vocalizations, the relationship between body weight and frequency content was similar for males and for females, suggesting that sex differences in call structure at adulthood are likely a by-product of sex differences in overall body size. Conversely, I found clear sex differences in the relationship between body size and the frequency content of alarm calls. Both measurements relating to frequency content underwent a

disproportionate drop in males as they grew relative to females, suggesting that the sex differences that characterize the spectral parameters of vervet monkey alarm calls are beyond that which is predicted by sex differences in body size. I also found strong sex differences in the effect of body size on the variability in the lowest dominant frequency band (DF1) in alarm calls; although DF1 became less variable in the alarm calls of both males and females as they grew, the calls produced by the largest males (adults) were less variable than the calls produced by the largest females. The first dominant frequency band of the alarm calls produced by males seemed to converge on a relatively narrow range of frequencies between 176 HZ and 826 HZ, while the same acoustic parameter for females remained highly variable in adults (174 Hz to 3043 Hz). One interpretation of this finding is that this parameter is under strong directional selection in the alarm calls produced by adult males. Importantly, these sex differences in the relationship between body size and spectral parameters were only found in alarm calls, suggesting that the extended low frequency content of alarm calls does not generalize to the entire vocal repertoire.

The logical next question would be to determine what possible selective forces have led to this sex-related divergence in alarm structure. Specifically, why do the spectral characteristics of the male's alarm call decrease so sharply as they grow? From an information based standpoint, the terrestrial alarms produced by males and females should both "reference" the same general predator classes; regardless of the acoustic variation that characterizes the calls of males and females, their tendency to be produced in association with predator encounters suggests that from a receiver's point of view, there should be little difference in what the signals "stand for". Why then do males produce alarm calls that are characterized by such low frequency acoustic energy compared to females?

Communication with the predator

Predator deterrence

One possibility is that the alarm calls generally produced by males are adapted to be transmitted over longer distances than those produced by females. Because lower frequency sounds attenuate less as they travel from their source (Bradbury and Vehrencamp 2011), the alarm calls produced by adult males are likely to propagate over a relatively large distance compared to those produced by females. It was often the case that I could clearly hear threat alarm barks being produced by neighboring troops in the field. Conversely, the female's "chirp" vocalization was comparatively harder to hear from a distance (personal observation). The increased range of male alarm calls could thus allow receivers who are further away from the caller to perceive the signal. Although research has suggested that vervet monkeys are less responsive to alarm calls produced by distant callers (Ducheminsky et al. 2014), it is possible that this extended range may facilitate communication with potential predators. While alarm calls are often assumed to function primarily as warning signals used to alert group members about the presence of a potential threat, there is also evidence that alarm calling may benefit signalers directly through communication with the predator itself (Caro 2005). For instance, the "perception advertisement" hypothesis suggests that prey species may produce signals that communicate to the predator that it has been detected. In instances where a prey animal detects a predator at a distance from which it is unlikely to be caught, it may be adaptive to make the predator aware that it has been detected, as predators are likely to give up their pursuit to avoid wasting energy on a chase that is likely to be energetically costly with little chance of pay off (Caro 2005).

The existence of perception advertisement signals have been proposed in both visual and auditory modes in a number of species (Eastern Swamphen, *Porphyrio melanotus*: Woodland et

al. 1980, Hares: Holley 1993; Klipspringer, *Oreotragus oreotragus*: Tilson and Norton 1981, Randall and Matocq 1977; Primates: Zuberbühler et al. 1999), and it has even been proposed that in rodents, alarm signals may have originally evolved for this purpose before being co-opted as warning signals (Shelley and Blumstein 2004). Importantly, there is good evidence to suggest that the terrestrial alarm calls produced by vervet monkeys function in predator deterrence in addition to functioning as a warning signal for conspecifics (Isbell and Bidner 2016). By tracking the movements of GPS-collared leopards, researchers found that predators within 200m of a group of alarm-calling vervet monkeys actively moved away, suggesting that perception advertisement is one possible function of calling. This being the case, the increased range of the male's terrestrial alarm call may allow these signals to be heard by predators well before they are close enough to callers to represent an immediate threat. One issue with this hypothesis, however, is that it does not account for the sex differences in alarm structure; functionally it seems likely that females could also benefit from deterring pursuit from predators, and it would be difficult to explain the apparent sex differences in call structure based on this hypothesis alone. Additionally, the anecdotal observations of Seyfarth et al. (1980) that females and juveniles are more likely to produce male-like calls when in close proximity to predators is in conflict with this hypothesis.

Aggression towards predators

Another, non-mutually exclusive explanation for the lower frequency content of the adult male alarm call is that they represent a form of aggressive signal directed towards predators. It was often the case during predator mount presentations that adult males would actively move towards the predator mount while calling. In fact, although the subjects never got close enough to the mount to make physical contact with it, it did seem in some instances that they were in a sense mobbing

the mount, in that they would move into the trees and shrubs closest to the mount, and call towards it until it was covered and removed (personal observation). Vocal signals produced in aggressive contexts are often characterized by energy distribution focused in the lower frequencies, and this may function to increase the perceived size of callers, thus acting as a form of intimidation towards receivers (Morton 1977). It is conceivable that a predator who has lost the element of surprise, and is surrounded by multiple males producing loud, ‘deep’ sounds may be dissuaded from engaging physically with any one caller for fear of physical injury. That the threat alarm bark is associated with an aggressive motivational state in callers is supported by the fact that the terrestrial alarm calls of adult males have been found to grade in structure with signals produced in aggressive contexts with other males (Price et al. 2015).

A related possibility is that the lower frequency content of the male alarm call functions as a form of “quality advertisement”, whereby producing sounds that emulate those of larger bodied individuals conveys a sense that a caller is larger and of a higher quality to a predator, and is therefore liable to escape if pursued (Caro 2005). These hypotheses are in line with the fact that vervet females can also produce signals that are acoustically somewhat more similar to the archetypal male “threat alarm bark”, and anecdotal evidence that suggests they tend to do so when they are closer to the predator that elicited calling (Seyfarth et al. 1980b). Thus, despite the fact that the calls analysed here were produced in response to similar external stimuli, it is possible that call structure may vary as a by-product of the caller’s motivational state. The tendency for males and females to produce calls with different acoustic properties may actually represent differences in the motivational responses of each sex to the detection of predators. Specifically, males may take an offensive, or aggressive approach to predators more commonly than females. Still, when females are in closer proximity to predators, or encounter smaller predators, they may be more

likely to produce the “bark-like” alarm call. Future research should explore the effects of the size and proximity of the eliciting stimulus on the acoustic structure of alarms produced by callers of both sexes.

Sexual selection

Another possibility is that the observed sex differences in the terrestrial alarm calls are the product of sexual selection. There is evidence in a number of species that females exhibit preference for vocalizations produced by larger males (Schwartz 1986, Forester and Czarnowsky 1985, Giacoma et al. 1997, Charlton et al. 2007b, 2012). Such preferences are likely to put a premium on acoustic properties that effectively exaggerate the body size of a signaler (Fitch 1999, Taylor and Reby 2010). The acoustic exaggeration hypothesis suggests that in some instances, there may be selection on callers to produce vocal signals which functionally exaggerate body size for the purpose of advertising resource holding potential to either male competitors, or potential mates (Fitch 1999).

With this in mind, it is possible that the acoustic structure of the male’s terrestrial alarm call is a by-product of intersexual selection for features that influence female mate choice. In line with this hypothesis, male vervet monkeys under experimental conditions produced alarm calls at higher rates when in the company of adult females as opposed to other males (Cheney and Seyfarth 1985b). One interpretation of this finding was that males may benefit from warning potential mates to the presence of a predator, as they may gain favor with those females. Predator encounters may represent a context in which males are able to display some aspect of quality or vigor by engaging in risky behaviours while producing conspicuous signals that may draw the attention of females

(Zuberbühler 2006; but see Wilson and Evans 2008 for a discussion of alarm calling in chickens as a form of mate investment). This may also explain my results which indicated that terrestrial alarm calls vary predictably in structure between different individuals (see **Chapter 3**); individual differences in call structure would allow females to monitor which males participate actively in antipredator defense, and would allow females to assess the qualities of specific males based on individually-specific variation in call structure (Snowdon 2004, Delgado 2006). In addition to the direct benefits of attracting female attention, male vervet monkeys who are highly integrated with females in terms of grooming and/or spatial associations seem to benefit through increased power trajectories (Young et al. 2017), thus suggesting that increased association with females may also influence their standings amongst other males.

Male alarm calls could also function in male-male competition more directly, with alarm calls giving potential rivals the ability to assess each other's competitive ability without the need to engage with one another physically (Bradbury and Vehrencamp, 1998; Owings and Morton, 1998, Fitch 2006). Recent work has shown that the male's threat alarm bark is acoustically similar to vocalizations that are produced in association with male-male aggression (Price et al. 2015), giving some credence to the possibility that these signals function as a sexually selected trait, which could be used by males to advertise competitive ability to other males during aggressive interactions. This has been proposed for the alarm calls of Diana monkeys (*Cercopithecus diana*), who also produce sexually dimorphic alarm calls, with males producing lower-pitched, louder signals relative to those of females (Zuberbühler et al. 1997). Because Diana monkeys live in single male groups, male-male competition over access to females is predicted to be high. As such, selection for secondary sexual characteristics is predicted to be high as well. Baboons also produce Wahoo vocalizations in association with both predatory contexts, as well as during male-male

contests (Fischer et al. 2002), suggesting that a relationship between alarm calls and intersexual selection may be widespread, at least among cercopithecine primates.

Of course, the ability for an individual to exaggerate their size effectively hinges on the absence of other, more reliable cues to body size. Because vervets typically live in multimale groups (Isbell et al. 2004), males living within the same group are unlikely to rely simply on vocal cues to assess each other's competitive abilities. However, male vervet monkeys have been reported to disperse at high rates during the mating season, and are able to gain access to mating opportunities upon entering a new social group (Henzi and Lucas 1980, Henzi 1982). Males have also been reported to engage in elevated levels of physical aggression with one another during these times, and incur increased numbers of wounds as a result (Henzi 1982, Freeman 2012). This suggests that males may benefit from discouraging other males from immigrating into their own group. Thus, it is possible that the threat alarm bark, in addition to functioning as an alarm call, may act to advertise a male's presence and competitive ability to extra-group males. This would also provide some explanation as to the increased range of these calls relative to those of females; because the alarm bark could be heard by males in neighboring groups, callers may dissuade extra-group males from immigration into their groups, thus reducing future competition over access to mating opportunities (Isbell et al. 2004). Playback experiments have shown that black howler monkeys seem to monitor the number of potential rival males present during intergroup encounters based on the number of individual callers in another social group (Kitchen 2004). Vervets may similarly assess the number of rival males in neighboring groups, and avoid migrating into groups with more competitors.

While the present study examined the effects of body weight on call structure throughout ontogeny, future studies should focus on correlations between call structure and specific qualities

of adult signalers, including body size, fighting ability, stamina, dominance rank, or even reproductive success (Delgado 2006).

Chapter 5 General Discussion

5.1 Overview

The overarching goal of my study was to describe possible sources of acoustic variation in some of the vocal signals produced by vervet monkeys. In particular, I focused my analysis on potential sources of variation in two call types – grunts and terrestrial alarm calls. My independent variables of interest included those intrinsic to individual callers, such as age, sex, body weight and identity. I also investigated whether grunts varied in structure based on the contexts in which they were produced. In my analyses, I considered the effects these variables have on both temporal and frequency characteristics of each respective call type. In what follows, I briefly summarize my findings. I then discuss their relevance for the study of primate vocal-communication as a whole, and discuss some unresolved issues that could provide the foundation for future research.

5.2 Summary of findings

Chapter 2: Acoustic variability in the structure of vervet grunts

In **Chapter 2**, I attempted to determine whether the grunts produced by vervet monkeys vary in structure based on the caller's age, sex, or identity. I also investigated the degree to which the acoustic structure of grunts predicted production context, using a broader range of grunts than have been used in previous studies. The most important empirical findings of this chapter were:

- 1. Grunts produced by vervet monkeys contain acoustic cues pertaining to both the age, and sex of the caller.**
- 2. Grunts produced by adults do not seem to vary in structure reliably between different callers, suggesting that identity cueing is not a primary function of this vocal signal.**
- 3. There was no relationship between call structure and the contexts in which grunts were emitted.**

The finding that vervet grunts varied in structure based on the caller's age and sex is consistent with the general trend across non-human primates: Age and sex differences in caller body size tend to correlate with age/sex differences in the structure of the vocalizations produced by different individuals (Ey et al. 2007a). Conversely, the lack of clear cues to caller identity was surprising, particularly in light of the fact that grunts are often produced in social contexts, where caller identity is likely to play an important role. One possible explanation for this is the fact that the close-range use of grunts may render acoustic cues to caller ID redundant as receivers are likely able to use visual or olfactory cues to identify callers in these instances (Marler 1967). The most surprising finding in this chapter was that, contrary to the findings of previous studies, vervet grunts did not seem to vary in predictable ways based on the contexts in which they were produced. Again, I discuss the possibility that there may be little selective pressure on grunts to provide clear cues to context, as they are generally produced in contexts where inter-individual visibility is good, allowing receivers to determine the context underlying their production through the visual channel.

Chapter 3: Acoustic variation in the structure of terrestrial alarm calls

In **Chapter 3**, I explored the degree to which the acoustic structure of terrestrial alarm calls in vervet monkeys varied based on the sex of the caller. I used a statistical clustering technique to determine whether the 'chirps' produced by adult females exist as distinct calls from the 'threat-alarm-barks' produced by males. I also explored whether the terrestrial alarm calls of vervet monkeys were individually recognizable. The key findings in the chapter were:

- 1. The calls of adult males are acoustically distinguishable from those produced by adult females.**
- 2. The acoustic variables that best distinguished between these calls were all related to the distribution of acoustic energy throughout the call, and included the frequency of**

the lowest dominant frequency band, the peak frequency of the call, and the overall distribution of energy across the entire call (DFA1, DFA2).

- 3. Although the calls of adult males and adult females are generally highly distinguishable, there seems to be acoustic intermediates between the female's chirp and the male's threat alarm bark.**
- 4. The alarm calls of both males and females vary in structure based on caller identity.**

The finding that males produced terrestrial alarm calls that could be systematically distinguished in their acoustic structure from those of adult females is consistent with previous subjective descriptions of the vervet's terrestrial alarm calls (Struhsaker 1967a). Still, this is the first time these sex differences have been assessed quantitatively. The cluster analysis highlighted the fact that these sex differences do not amount to two clearly distinct call types, however, as the female "chirp" and the male "threat alarm bark" are characterized by some intergradation. Specifically, males produced some calls that clustered with the stereotypical female chirp. Females also produced graded intermediates between both call types. Finally, my analysis of individual distinctiveness in call structure revealed that alarm calls do vary in structure among individuals. I discuss the possible implications for these identity cues for the function of alarm calls as warning signals.

Chapter 4: The influence of body size on the acoustic structure of grunts and terrestrial alarm calls

In **Chapter 4**, I attempted to determine whether sex differences in the structure of grunts and terrestrial alarm calls were a by-product of sex differences in body size. Specifically, I used body weight as a proxy for body size, and used a series of linear models to determine whether the relationship between body weight and each acoustic variable was consistent for males and females.

The key findings in this chapter were:

1. **The relationship between the acoustic structure of grunts and the caller's body weight was similar for both males and females, suggesting that sex differences in the structure of grunts in adults are likely to be the by-product of sex differences in body size.**
2. **The relationship between the temporal parameters of terrestrial alarm calls and the caller's body weight was similar for both males and females, suggesting that sex differences in adults are likely to be the by-product of sex differences in body size.**
3. **The relationship between the spectral parameters of terrestrial alarm calls and body weight differed for males and females. Although both measurements of energy distribution decreased as callers grew in size, this trend was more pronounced in males than in females.**
4. **Variability in the lowest dominant frequency band also decreased disproportionately more in males relative to females as they matured.**

The results of this chapter suggest that there are asymmetries in the extent of sex differences in vocal signals of vervet monkeys. Specifically, while the sex differences that characterize the grunts of adult vervet monkeys seem to be linked to differences in body size (and by extension, sex differences in the size of the components of the vocal tract), sex differences in the spectral parameters of their terrestrial alarm calls are disproportionate to the sex differences in body size. I discuss a number of possible explanations for the sex differences in the vervet monkey's terrestrial alarm calls. For instance, these differences could reflect different motivations underlying the production of calls by males and females, or could alternatively be the result of sexual selection on the loud (alarm) calls of adult males.

5.3 Implications for the field of primate communication

Functional reference and “arbitrary” signal structure

The last four decades of research into the vocal signaling behaviour of non-human primates have been dominated by a focus on functional reference. This focus is understandable, as some researchers view these functionally-referential signals as representing a possible continuity

between the vocal utterances of non-human animals and human language (Scarantino and Clay 2015). Still, some authors have argued that this linguistically-based interpretation has diverted attention away from more classical biological constructs; specifically, the supposition that the form and function of vocal signals are inter-connected (Owren and Rendall 2001, Owings and Morton 1998, Morton 2017).

Taken together with the body of research amassed to date, my own research further highlights the point that the relationship between signal structure, function and mechanism is a complicated one, in which selection is likely to result in call structures that vary along multiple dimensions simultaneously. In the case of acoustic signals, structural variation is invariably constrained and shaped by anatomical variation in individual callers, and so it is possible that receiver response will involve integrating, at the very least, cues to intrinsic characteristics of the signaller, as well as potentially those associated with context.

In **Chapter 3** I show that in addition to their purported function as predator specific warning signals (Seyfarth et al. 1980a,b), terrestrial alarm calls contain reliable cues to both caller sex and identity. Additionally, the results presented in **Chapter 4** suggest that the relationship between body size and the spectral parameters of alarm calls are actually different for males and females, indicating that sex differences in alarm structure are not merely the product of sex differences in body size. In many ways, it is surprising that these sex differences have not been investigated until now; the terrestrial alarm calls of vervet monkeys were subjectively classified into two sex-specific call “types” as far back as the original descriptions provided by Struhsaker in 1967 (Struhsaker 1967a). Still, sex differences in the alarm calling behaviour of vervet monkeys has received little empirical attention, which has left a conceptual gap in our knowledge concerning the different selective pressures acting on the signaling behaviour of males and females. My

research highlights the fact that functional reference only represents one aspect of the diverse mechanisms that underlie signal function, and hopefully stands as a starting point for future research into the function of these signals.

Future research should focus on possible selective pressures that may underlie this acoustic variation. Importantly, the acoustic parameters that seem to best distinguish between the alarms of males and females may provide some hints as to the possible function of these sex differences. The acoustic parameters that characterize male alarm calls for instance share some properties with long calls produced in a number of other species. In most species, long calls are produced by males exclusively, and are considered inter-group signals, meaning they are used to communicate between dispersed social groups (Mitani and Stuht 1998, Wich and Nunn 2002). As such, they are often characterized by high amplitudes and low frequencies, making them less susceptible to the affects of attenuation, thus allowing them to travel relatively long distances. From a functional point of view, long calls may serve a number of functions, including mate attraction, mate defence, and resource defence (reviewed in Delgado 2006). Although it seems clear that one function of the vervet male alarm bark is to act as a warning signal (Seyfarth et al. 1980a,b, but see Ducheminsky et al. 2014), the observed sexual dimorphism in call structure suggests that alarm calls are also influenced by sexual selection. As discussed elsewhere (**Chapter 4**), some researchers suggest that alarm calls primarily function as aggressive signals towards predators, and that their function as warning signals evolved only secondarily (Shelley and Blumstein 2004). The male vervet threat-alarm-bark may represent another instance in which a signal that evolved to serve one function has been additionally shaped by selection to serve a secondary purpose. Although my study does not address this hypothesis directly, it does highlight the fact that a focus on signal structure is likely

to generate hypotheses as to the possible functions of vocal signals that were left unexplored under purely referential paradigms.

The informational content of signals

As reviewed in **Chapter 1**, information-centered views suggest that whenever there is a predictable relationship between the acoustic features of a signal and some aspect of the signaler or the environment, receivers are able to make inferences (i.e., are informed) about those features upon perceiving the signal (Seyfarth et al. 2010). Of course, the results of my study could easily be interpreted under this information-based view. For instance, terrestrial alarm calls could be interpreted as carrying information about the presence of a predator, the identity of the caller, and the caller's sex. Grunts on the other hand could be interpreted as carrying information regarding caller age, and sex, but not caller identity or context. As discussed in **Chapter 2**, it is possible that grunts are not variable enough to convey information pertaining to caller ID or context because receivers are able to gain this information through other channels. In contrast with alarm calls, grunts are relatively low in amplitude, suggesting that most receivers are likely in relatively close proximity to callers when the signal is received. This would allow receivers to visually assess caller identity, and to visually determine the context that elicited calling. This being the case, it seems likely that receivers could also assess caller age and sex in the same way, suggesting selection should be equally as weak for grunts to encode this information as well. Under this interpretation, age- and sex-related cues in grunts may simply be the inevitable by-product of broad size differences between different age/sex classes, meaning it is possible that they have not been

selected to convey this information per se. This of course leads to the obvious question: What is the function of grunt vocalizations?

One possibility is that the function of grunts is not to convey any “information” to receivers at all. As highlighted by Owren and Rendall (2001), vocal signals can exert their effects on receivers without the need for any information transfer to have occurred. Grunts could, for instance, simply act as an attention-grabbing sound that could function to reorient a receiver’s attention towards the caller, or simply alert the receiver that another individual is nearby. Owren et al. (2005) also highlight the fact that grunts in a number of primate species tend to be produced in association with relatively innocuous contexts. As such, over repeated interactions, it is likely that receivers would begin to form associations between these calls and benign social outcomes. This being the case, grunts may simply elicit learned affects from receivers, thus functionally decreasing any potential uncertainty associated with close range interactions. Similarly, it is possible that grunts emitted during aggressive interactions may function to pacify attackers by stimulating a more relaxed affective state in receivers. Future studies should consider these possibilities, and continue to explore non-information centered mechanisms by which close range vocalizations in vervet monkeys, as well as in other species, exert their effects on receivers. For instance, researchers should determine whether the production of grunts as individuals approach one another has obvious effects on the quality of their subsequent interactions.

5.4 Limitations of my study, and directions for Future research

One limitation of my study is that it focuses exclusively on signal production. Communication involves both signalers and receivers, and it is critical that future research focuses

on determining whether the individually-specific cues highlighted by my analyses are biologically relevant and salient to receivers. Playback experiments, for example could be used to determine whether age, sex, or size differences in the structure of grunts or alarm calls affect receiver responses. Rendall et al. (2004) were able to show that baboons could distinguish between the grunts of males and females using laboratory-trained animals. Subjects were trained to discriminate between the calls of a representative male and female using an operant conditioning paradigm. This training subsequently generalized to grunts produced by other males and females, suggesting receivers could distinguish between calls based on the sex of the caller.

Future analyses should also explore functional explanations for the observed sex differences in the vervet monkey's terrestrial alarm call. Playback experiments could be used to test the responses of conspecifics to both the "chirp", and the "threat alarm bark", as well as to the acoustic intermediates that exist between these two call "types". Additionally, the effectiveness of these calls in deterring predators should be tested. Previous studies have explored the predator deterrence function of alarm calls using radio-collared predators, and tracking their movements after naturally occurring bouts of alarm calls (e.g., Zuberbühler et al 1999, Isbell and Bidner 2016). To elucidate the effects of calls produced by males and females respectively, future studies could use a playback design whereby the calls of either males or females are played to collared predators, thus allowing the effects of caller sex to be investigated under controlled circumstances. Experimenters could also vary the distance from which the playbacks are emitted in an attempt to determine whether the acoustic properties of the male's call allow them to ward off predators from a greater distance than the calls produced by females (see Discussion, **Chapter 4**).

Research could also focus on the possibility that sex differences in the terrestrial alarm calls of vervet monkeys represent a sexually selected trait. Snowdon (2004) pointed out that vocal

signals must meet a series of criteria to be considered the product of sexual selection. My results suggest that terrestrial alarm calls meet two of these criteria; specifically, my results indicate that terrestrial alarm calls are a) sexually dimorphic, and b) vary in structure between individuals, thus providing a substrate on which selection could act (**Chapter 3**). Still, future analyses should focus on whether variation in call structure is at all related to caller specific characteristics such as dominance status, or reproductive output. As primate loud calls have also been suggested to serve a spacing function in many instances (Mitani and Stuht 1998, Oliveira and Ades 2004, Delgado 2006), future research could also monitor the responses of neighboring groups, or extra group males to the vervet monkey's terrestrial alarm call. Finally, it remains to be determined whether receivers show some preference (inter-sexual selection), or aversion (intra-sexual selection) for signals based on variation between the calls of different males.

Future analyses could also attempt to explore the physiological underpinnings of the observed differences in the structure of grunts and alarm calls. The results of **Chapter 4** seem to indicate that the spectral parameters of the terrestrial alarm calls of males have decoupled from body size. Future studies should focus on the precise mechanisms by which male vervet monkeys produce their low frequency alarm calls. Radiographs or excised vocal tracts could be used to measure the various components of the male and female's vocal tract (Riede and Fitch 1999, Riede et al. 2005, Plotsky et al. 2013). Using these data, it may be possible to determine whether differences in vocal structure relate to sex-based differences in vocal anatomy alone. Fitch (2000) used x-ray video recordings of dogs (*Canis familiaris*), goats (*Capra hircus*), pigs (*Sus scrofa*) and cotton-top tamarins (*Saguinus oedipus*) to monitor the vocal tract configuration of live animals during phonation. Using this method, the acoustic structure of alarm calls could be better

understood in terms of differences in the position of the larynx and the articulators of males and females while producing calls.

Such a physiological approach to the analysis of signal structure may also reveal additional variation that was not detected by my analyses. For instance, future research should attempt to determine whether differences in the size and shape of the vocal tracts of different individuals correspond to specific resonant frequencies within the calls they produce (i.e., formants). While my analysis did not find clear individual differences in call structure within grunts, my frequency-based measurements centered around the overall distribution of energy within the call, and did not focus specifically on vocal resonances. Formants have been shown to act as reliable cues to caller identity in a number of species (e.g., rhesus monkey; Rendall et al. 1998, African elephant; Soltis et al. 2005, Fallow deer; Vannoni and McElligott 2007, Baboons; Rendall 2003, red-bellied lemurs; Gamba et al. 2012, Giant panda; Charlton et al. 2009). While the distribution of frequency amplitudes used in my analyses provides a useful way to quantify the distribution of energy throughout the call (Fischer et al. 2013), it is possible that a more focused formant analysis could reveal individual differences in call structure that were not detected by my analysis.

References

- Aisenberg, A. (2009). Male performance and body size affect female re-mating occurrence in the orb-web spider *Leucauge mariana* (Araneae, Tetragnathidae). *Ethology*, 115(12), 1127-1136.
- Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, 49(3-4), 227-266.
- Arnedo, L. F., Mendes, F. D., and Strier, K. B. (2010). Sex differences in vocal patterns in the northern muriqui (*Brachyteles hypoxanthus*). *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 72(2), 122-128.
- Bachorowski, J. A., and Owren, M. J. (1995a). Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychological Science*, 6(4), 219-224.
- Bachorowski, J. A., and Owren, M. J. (1995b). Vocal expression of emotion is associated with formant characteristics. *The Journal of the Acoustical Society of America*, 98(5), 2936-2937.
- Balcombe, J. P. (1990). Vocal recognition of pups by mother Mexican free-tailed bats, *Tadarida brasiliensis mexicana*. *Animal Behaviour*, 39(5), 960-966.
- Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, 70(3), 614.
- Barnard, C. J., and Burk, T. (1979). Dominance hierarchies and the evolution of "individual recognition". *Journal of Theoretical Biology*, 81(1), 65-73.
- Bauers, K. A. (1993). A functional analysis of staccato grunt vocalizations in the stump-tailed macaque (*Macaca arctoides*). *Ethology*, 94(2), 147-161.
- Bee, M. A., Perrill, S. A., & Owen, P. C. (1999). Size assessment in simulated territorial encounters between male green frogs (*Rana clamitans*). *Behavioral Ecology and Sociobiology*, 45(3-4), 177-184.
- Bergman, T. J., Beehner, J. C., Cheney, D. L., and Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science*, 302(5648), 1234-1236.
- Bernstein, I. S., and Ehardt, C. L. (1985). Agonistic aiding: kinship, rank, age, and sex influences. *American Journal of Primatology*, 8(1), 37-52.
- Blanckenhorn, W. U., Mühlhäuser, C., Morf, C., Reusch, T., and Reuter, M. (2000). Female choice, female reluctance to mate and sexual selection on body size in the dung fly *Sepsis cynipsea*. *Ethology*, 106(7), 577-593.

- Blumstein, D. T. (1999). The evolution of functionally referential alarm communication: multiple adaptations; multiple constraints. *Evolution of Communication*, 3(2), 135-147.
- Blumstein, D. T., and Daniel, J. C. (2004). Yellow-bellied marmots discriminate between the alarm calls of individuals and are more responsive to calls from juveniles. *Animal Behaviour*, 68(6), 1257-1265.
- Blumstein, D. T., Verneyre, L., and Daniel, J. C. (2004). Reliability and the adaptive utility of discrimination among alarm callers. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1550), 1851-1857.
- Boncoraglio, G., and Saino, N. (2007). Habitat structure and the evolution of bird song: a meta-analysis of the evidence for the acoustic adaptation hypothesis. *Functional Ecology*, 21(1), 134-142.
- Bouchet, H., Blois-Heulin, C., Pellier, A. S., Zuberbühler, K., and Lemasson, A. (2012). Acoustic variability and individual distinctiveness in the vocal repertoire of red-capped mangabeys (*Cercocebus torquatus*). *Journal of Comparative Psychology*, 126(1), 45.
- Bradbury, J. W., and Vehrencamp, S. L. (2011). *Principles of Animal Communication*. (2nd ed.). Sunderland, Massachusetts: Sinauer.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology*, 288(1), 1-20.
- Bugnyar, T., Kijne, M., and Kotrschal, K. (2001). Food calling in ravens: are yells referential signals? *Animal Behaviour*, 61(5), 949-958.
- Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1-28.
- Caine, N. G., and Stevens, C. (1990). Evidence for a “monitoring call” in red-bellied tamarins. *American Journal of Primatology*, 22(4), 251-262.
- Caro, T. (2005). *Antipredator defenses in birds and mammals*. University of Chicago Press.
- Cäsar, C., and Zuberbühler, K. (2012). Referential alarm calling behaviour in New World primates. *Current Zoology*, 58(5), 680-697.
- Chapman, C. A., Fedigan, L. M., and Fedigan, L. (1988). Ecological and demographic influences on the pattern of association in St. Kitts vervets. *Primates*, 29(3), 417-421.
- Chapman, C., and Fedigan, L. M. (1984). Territoriality in the St Kitts vervet, *Cercopithecus aethiops*. *Journal of Human Evolution*, 13(8), 677-686.

- Charlton, B. D., Ellis, W. A., Brumm, J., Nilsson, K., and Fitch, W. T. (2012). Female koalas prefer bellows in which lower formants indicate larger males. *Animal Behaviour*, *84*(6), 1565-1571.
- Charlton, B. D., Reby, D., and McComb, K. (2007). Female perception of size-related formant shifts in red deer, *Cervus elaphus*. *Animal Behaviour*, *74*(4), 707-714.
- Charlton, B. D., Reby, D., and McComb, K. (2007). Female red deer prefer the roars of larger males. *Biology Letters*, *3*(4), 382-385.
- Charlton, B. D., Zhihe, Z., and Snyder, R. J. (2009). The information content of giant panda, *Ailuropoda melanoleuca*, bleats: Acoustic cues to sex, age and size. *Animal Behaviour*, *78*(4), 893e898.
- Charrad, Nadia Ghazzali, Veronique Boiteau, Azam Niknafs (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, *61*(6), 1-36. URL <http://www.jstatsoft.org/v61/i06/>.
- Charrier, I., Ahonen, H., and Harcourt, R. G. (2011). What makes an Australian sea lion (*Neophoca cinerea*) male's bark threatening?. *Journal of Comparative Psychology*, *125*(4), 385.
- Charrier, I., Jouventin, P., Mathevon, N., and Aubin, T. (2001). Individual identity coding depends on call type in the South Polar skua *Catharacta maccormicki*. *Polar Biology*, *24*(5), 378-382.
- Charrier, I., Mathevon, N., and 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., and Jouventin, P. (2003). Vocal signature recognition of mothers by fur seal pups. *Animal Behaviour*, *65*(3), 543-550.
- Chen, H. C., Kaplan, G., and Rogers, L. J. (2009). Contact calls of common marmosets (*Callithrix jacchus*): influence of age of caller on antiphonal calling and other vocal responses. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, *71*(2), 165-170.
- Cheney, D. L. (1977). The acquisition of rank and the development of reciprocal alliances among free-ranging immature baboons. *Behavioral Ecology and Sociobiology*, *2*(3), 303-318.
- Cheney, D. L., and Seyfarth, R. M. (1980). Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour*, *28*(2), 362-367.
- Cheney, D. L., and Seyfarth, R. M. (1981). Selective forces affecting the predator alarm calls of vervet monkeys. *Behaviour*, *76*(1-2), 25-60.
- Cheney, D. L., and Seyfarth, R. M. (1982a). How vervet monkeys perceive their grunts: field playback experiments. *Animal Behaviour*, *30*(3), 739-751.

- Cheney, D. L., and Seyfarth, R. M. (1982b). Recognition of individuals within and between groups of free-ranging vervet monkeys. *American Zoologist*, 22(3), 519-529.
- Cheney, D. L., and Seyfarth, R. M. (1983). Nonrandom dispersal in free-ranging vervet monkeys: social and genetic consequences. *The American Naturalist*, 122(3), 392-412.
- Cheney, D. L., and Seyfarth, R. M. (1985a). Social and non-social knowledge in vervet monkeys. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 308(1135), 187-201.
- Cheney, D. L., and Seyfarth, R. M. (1985b). Vervet monkey alarm calls: Manipulation through shared information? *Behaviour*, 94(1-2), 150-166.
- Cheney, D. L., and Seyfarth, R. M. (1986). The recognition of social alliances by vervet monkeys. *Animal Behaviour*, 34(6), 1722-1731.
- Cheney, D. L., and Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour*, 36(2), 477-486.
- Cheney, D. L., and Seyfarth, R. M. (1989). Redirected aggression and reconciliation among vervet monkeys, *Cercopithecus aethiops*. *Behaviour*, 258-275.
- Cheney, D., and Seyfarth, R. (1990a). Attending to behaviour versus attending to knowledge: Examining monkeys' attribution of mental states. *Animal Behaviour*, 40(4), 742-753.
- Cheney, D. L., and Seyfarth, R. M. (1990b). *How monkeys see the world: Inside the mind of another species*. University of Chicago Press.
- Cheney, D. L., and Seyfarth, R. M. (1997). Reconciliatory grunts by dominant female baboons influence victims' behaviour. *Animal behaviour*, 54(2), 409-418.
- Cheney, D. L., and Seyfarth, R. M. (1998). Why animals don't have language. *Tanner lectures on human values*, 19, 173-210.
- Cheney, D. L., and Seyfarth, R. M. (2005). Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review*, 22(2-4), 135-159.
- Cheney, D. L., and Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, 115(9), 1974-1979.
- Cheney, D. L., Seyfarth, R. M., and Silk, J. B. (1995). The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: evidence for causal reasoning? *Journal of Comparative Psychology*, 109(2), 134.
- Clarke, E., Reichard, U. H., and Zuberbühler, K. (2015). Context-specific close-range “hoo” calls in wild gibbons (*Hylobates lar*). *BMC evolutionary biology*, 15(1), 56.

- Clay, Z., and Zuberbühler, K. (2009). Food-associated calling sequences in bonobos. *Animal Behaviour*, 77(6), 1387-1396.
- Clay, Z., and Zuberbühler, K. (2011). Bonobos extract meaning from call sequences. *PLoS One*, 6(4), e18786.
- Clutton-Brock, T. H., Albon, S. D., Gibson, R. M., and Guinness, F. E. (1979). The logical stag: adaptive aspects of fighting in red deer (*Cervus elaphus L.*). *Animal behaviour*, 27, 211-225.
- Collins, S. A. (2000). Men's voices and women's choices. *Animal behaviour*, 60(6), 773-780.
- Compton, L. A., Clarke, J. A., Seidensticker, J., and Ingrisano, D. R. (2001). Acoustic characteristics of white-nosed coati vocalizations: a test of motivation-structural rules. *Journal of Mammalogy*, 82(4), 1054-1058.
- Coss, R. G., McCowan, B., and Ramakrishnan, U. (2007). Threat-related acoustical differences in alarm calls by wild bonnet macaques (*Macaca radiata*) elicited by python and leopard models. *Ethology*, 113(4), 352-367.
- Crockford, C., Gruber, T., and Zuberbühler, K. (2018). Chimpanzee quiet hoo variants differ according to context. *Royal Society open science*, 5(5), 172066.
- Crockford, C., Herbinger, I., Vigilant, L., and Boesch, C. (2004). Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology*, 110(3), 221-243.
- Crockford, C., Wittig, R. M., and Zuberbühler, K. (2015). An intentional vocalization draws others' attention: a playback experiment with wild chimpanzees. *Animal cognition*, 18(3), 581-591.
- Crockford, C., Wittig, R. M., and Zuberbühler, K. (2017). Vocalizing in chimpanzees is influenced by social-cognitive processes. *Science advances*, 3(11), e1701742.
- Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., and Urban, J. (2002). Bioacoustics: Only male fin whales sing loud songs. *Nature*, 417(6891), 809.
- Darwin, C. (1872). *The expression of emotions in animals and man*. London: Murray.
- Darwin, C. (1896). *The descent of man and selection in relation to sex* (Vol. 1). D. Appleton.
- Davis, M. S. (1987). Acoustically mediated neighbor recognition in the North American bullfrog, *Rana catesbeiana*. *Behavioral Ecology and Sociobiology*, 21(3), 185-190.
- Dawkins, R., and Krebs, J. R. (1978). Animal signals: information or manipulation. *Behavioural ecology: An evolutionary approach*, 2, 282-309.
- Delgado, R. A. (2006). Sexual selection in the loud calls of male primates: signal content and function. *International Journal of Primatology*, 27(1), 5-25.

- Di Bitetti, M. S. (2003). Food-associated calls of tufted capuchin monkeys (*Cebus apella nigrinus*) are functionally referential signals. *Behaviour*, 140(5): 565-592.
- Dubreuil, C., Notman, H., and Pavelka, M. S. (2015). Sex differences in the use of whinny vocalizations in spider monkeys (*Ateles geoffroyi*). *International Journal of Primatology*, 36(2), 412-428.
- Ducheminsky, N., Henzi, S. P., and Barrett, L. (2014). Responses of vervet monkeys in large troops to terrestrial and aerial predator alarm calls. *Behavioral ecology*, 25(6), 1474-1484.
- Dunlop, R. A. (2017). Potential motivational information encoded within humpback whale non-song vocal sounds. *The Journal of the Acoustical Society of America*, 141(3), 2204-2213.
- Egnor, S. R., and Hauser, M. D. (2004). A paradox in the evolution of primate vocal learning. *Trends in neurosciences*, 27(11), 649-654.
- Evans, C. S. (1997). Referential signals. In *Communication* (pp. 99-143). Springer, Boston, MA.
- Evans, C. S., and Evans, L. (1999). Chicken food calls are functionally referential. *Animal behaviour*, 58(2), 307-319.
- Evans, C. S., Evans, L., and Marler, P. (1993). On the meaning of alarm calls: functional reference in an avian vocal system. *Animal Behaviour*, 46(1), 23-38.
- Evans, C. S., and Marler, P. (1994). Food calling and audience effects in male chickens, *Gallus gallus*: their relationships to food availability, courtship and social facilitation. *Animal Behaviour*, 47(5), 1159-1170.
- Ey, E., Hammerschmidt, K., Seyfarth, R. M., and Fischer, J. (2007a). Age- and sex-related variations in clear calls of *Papio ursinus*. *International Journal of Primatology*, 28(4), 947-960.
- Ey, E., Pfefferle, D., and Fischer, J. (2007b). Do age- and sex-related variations reliably reflect body size in non-human primate vocalizations? A review. *Primates*, 48(4), 253-267.
- Falls, J. B. (1982). Individual recognition by sounds in birds. *Acoustic communication in birds*, 2, 237-278.
- Faragó, T., Pongrácz, P., Miklósi, Á., Huber, L., Virányi, Z., and Range, F. (2010). Dogs' expectation about signalers' body size by virtue of their growls. *PLoS One*, 5(12), e15175.
- Faragó, T., Pongrácz, P., Range, F., Virányi, Z., and Miklósi, Á. (2010). 'The bone is mine': affective and referential aspects of dog growls. *Animal Behaviour*, 79(4), 917-925.
- Fedigan, L., and Fedigan, L. M. (1988). *Cercopithecus aethiops*: a review of field studies. In A. Gautier-Hion, R. Fourliere, J.-P. Gautier and J. Kingdon, (Eds) *A Primate Radiation: Evolutionary Biology of the African Guenons* (pp. 389-411). Cambridge: Cambridge University Press.

- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., and Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal behaviour*, 69(3), 561-568.
- Fichtel, C., and Kappeler, P. M. (2002). Anti-predator behavior of group-living Malagasy primates: mixed evidence for a referential alarm call system. *Behavioral Ecology and Sociobiology*, 51(3), 262-275.
- Fischer, J. (2017). Primate vocal production and the riddle of language evolution. *Psychonomic bulletin and review*, 24(1), 72-78.
- Fischer, J., Hammerschmidt, K., and Todt, D. (1998). Local variation in Barbary macaque shrill barks. *Animal Behaviour*, 56(3), 623e629
- Fischer, J., Hammerschmidt, K., Cheney, D. L., and 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.
- Fischer, J., Hammerschmidt, K., Cheney, D. L., and Seyfarth, R. M. (2001a). Acoustic features of female chacma baboon barks. *Ethology*, 107(1), 33-54.
- Fischer, J., Kitchen, D. M., Seyfarth, R. M., and Cheney, D. L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology*, 56(2), 140-148.
- Fischer, J., Metz, M., Cheney, D. L., and Seyfarth, R. M. (2001b). Baboon responses to graded bark variants. *Animal Behaviour*, 61(5), 925-931.
- Fischer, J., Noser, R., and Hammerschmidt, K. (2013). Bioacoustic field research: a primer to acoustic analyses and playback experiments with primates. *American journal of primatology*, 75(7), 643-663.
- Fischer, J., Wadewitz, P., and Hammerschmidt, K. (2017). Structural variability and communicative complexity in acoustic communication. *Animal behaviour*, 134, 229-237.
- Fisher, J. B. (1954). Evolution and bird sociality. In J. Huxley, A. C. Hardy and E. B. Ford (Eds.), *Evolution as a Process* (pp. 71-83). London: Allen and Unwin.
- 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.
- Fitch, W. T. (1999). Acoustic exaggeration of size in birds via tracheal elongation: comparative and theoretical analyses. *Journal of Zoology*, 248(1), 31-48.
- Fitch, W. T. (2000). The phonetic potential of nonhuman vocal tracts: comparative cineradiographic observations of vocalizing animals. *Phonetica*, 57(2-4), 205-218.

- Fitch, W. T. (2006). Production of vocalizations in mammals. *Visual Communication*, 3(2006), 145.
- Fitch, W. T., and 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.
- Fitch, W. T., and Hauser, M. D. (1995). Vocal production in nonhuman primates: acoustics, physiology, and functional constraints on “honest” advertisement. *American Journal of Primatology*, 37(3), 191-219.
- Fitch, W. T., and Hauser, M. D. (2003). Unpacking “honesty”: vertebrate vocal production and the evolution of acoustic signals. In A. Simmons, R. R. Fay, & A. N. Popper (Eds.), *Acoustic communication* (pp. 65-137). Springer, New York, NY.
- Fitch, W. T., and 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.
- Flombaum, J. I., and Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology*, 15(5), 447-452.
- Font, E., and Carazo, P. (2010). Animals in translation: why there is meaning (but probably no message) in animal communication. *Animal Behaviour*, 80(2), e1.
- Forester, D. C., and Czarnowsky, R. (1985). Sexual selection in the spring peeper, *Hyla crucifer* (Amphibia, Anura): role of the advertisement call. *Behaviour*, 92(1-2), 112-127.
- Fox, E. A. (2002). Female tactics to reduce sexual harassment in the Sumatran orangutan (*Pongo pygmaeus abelii*). *Behavioral Ecology and Sociobiology*, 52(2), 93-101.
- Freeman, N. J. (2012). *Some aspects of male vervet monkey behaviour* (Doctoral dissertation, Lethbridge, Alta.: University of Lethbridge, c2012).
- Fugate, J. M., Gouzoules, H., and Nygaard, L. C. (2008). Recognition of rhesus macaque (*Macaca mulatta*) noisy screams: evidence from conspecifics and human listeners. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 70(6), 594-604.
- Gall, G. E., and Manser, M. B. (2017). Group cohesion in foraging meerkats: follow the moving ‘vocal hot spot’. *Royal Society open science*, 4(4), 170004.
- Gamba, M., Colombo, C., and Giacoma, C. (2012). Acoustic cues to caller identity in lemurs: a case study. *Journal of Ethology*, 30(1), 191-196.
- Garcia, M., and Favaro, L. (2017). Animal vocal communication: function, structures, and production mechanisms. *Current Zoology*, 63(4), 417-419.
- Gautier, J. P., and Gautier-Hion, A. (1977). Communication in old world monkeys. In T. A. Sebeok (Ed.), *How Animals Communicate*. (pp. 890-964). Indiana University Press.

- Gelman, A., and Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, 9(6), 641-651.
- Giacoma, C., Zugolaro, C., and Beani, L. (1997). The advertisement calls of the green toad (*Bufo viridis*): variability and role in mate choice. *Herpetologica*, 53(4) 454-464.
- Gill, S. A., and Bierema, A. M. K. (2013). On the meaning of alarm calls: a review of functional reference in avian alarm calling. *Ethology*, 119(6), 449-461.
- Gill, S. A., and Sealy, S. G. (2004). Functional reference in an alarm signal given during nest defence: set calls of yellow warblers denote brood-parasitic brown-headed cowbirds. *Behavioral Ecology and Sociobiology*, 56(1), 71-80.
- Given, M. F. (1987). Vocalizations and acoustic interactions of the carpenter frog, *Rana virgatipes*. *Herpetologica*, 43(4), 467-481.
- Godard, R. (1991). Long-term memory of individual neighbours in a migratory songbird. *Nature*, 350(6315), 228.
- Goodall, J. (1985). *The Chimpanzees of Gombe*. Harvard University Press. Cambridge, MA.
- Goudbeek, M., and Scherer, K. (2010). Beyond arousal: Valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America*, 128(3), 1322-1336.
- Gouzoules, S., Gouzoules, H., and Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) screams: representational signalling in the recruitment of agonistic aid. *Animal behaviour*, 32(1), 182-193.
- Gouzoules, H., Gouzoules, S., and Marler, P. (1985). External reference and affective signaling in mammalian vocal communication. In R. G. Rawlins and M. J Kessler (Eds.), *The development of expressive behavior* (pp. 77-101). Academic Press.
- Gouzoules, H., Gouzoules, S., and Marler, P. (1986). Vocal communication: a vehicle for the study of social relationships. *The Cayo Santiago Macaques: History, Behavior and Biology* (pp. 111-129). State University of New York Press.
- Gouzoules, H., Gust, D. A., Donaghey, B., and Andre, E. S. (1998). Estrus vocalizations in two primate species (*Cercocebus torquatus atys* and *Macaca nemestrina*): evidence for an effect of intrasexual competition. *Evolution of Communication*, 2(2), 189-215.
- Gros-Louis, J. J., Perry, S. E., Fichtel, C., Wikberg, E., Gilkenson, H., Wofsy, S., and Fuentes, A. (2008). Vocal repertoire of *Cebus capucinus*: acoustic structure, context, and usage. *International Journal of Primatology*, 29(3), 641-670.
- Gruber, T., and Zuberbühler, K. (2013). Vocal recruitment for joint travel in wild chimpanzees. *PLoS One*, 8(9), e76073.
- Gustison, M. L., Johnson, E. T., Beehner, J. C., and Bergman, T. J. (2019). The social functions of complex vocal sequences in wild geladas. *Behavioral ecology and sociobiology*, 73(1), 14.

- Hammerschmidt, K., and Fischer, J. (1998). The vocal repertoire of Barbary macaques: a quantitative analysis of a graded signal system. *Ethology*, *104*(3), 203-216.
- Hammerschmidt, K., and Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice*, *21*(5), 531-540.
- Hammerschmidt, K., Newman, J. D., Champoux, M., and Suomi, S. J. (2000). Changes in rhesus macaque 'coo' vocalizations during early development. *Ethology*, *106*(10), 873-886.
- Hanson, M. T., and Coss, R. G. (2001). Age differences in the response of California ground squirrels (*Spermophilus beecheyi*) to conspecific alarm calls. *Ethology*, *107*(3), 259-275.
- Hare, B., Call, J., Agnetta, B., and Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, *59*(4), 771-785.
- Hare, B., Call, J., and Tomasello, M. (2001). Do chimpanzees know what conspecifics know?. *Animal behaviour*, *61*(1), 139-151.
- Hare, B., Call, J., and Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition*, *101*(3), 495-514.
- Hare, J. F., and Atkins, B. A. (2001). The squirrel that cried wolf: reliability detection by juvenile Richardson's ground squirrels (*Spermophilus richardsonii*). *Behavioral Ecology and Sociobiology*, *51*(1), 108-112.
- Harries, M., Hawkins, S., Hacking, J., and Hughes, I. (1998). Changes in the male voice at puberty: vocal fold length and its relationship to the fundamental frequency of the voice. *The Journal of Laryngology and Otology*, *112*(5), 451-454.
- Harris, T. R., Fitch, W. T., Goldstein, L. M., and Fashing, P. J. (2006). Black and white colobus monkey (*Colobus guereza*) roars as a source of both honest and exaggerated information about body mass. *Ethology*, *112*(9), 911-920.
- Harrison, M. J. (1983). Patterns of range use by the green monkey, *Cercopithecus sabaues*, at Mt. Assirik, Senegal. *Folia Primatologica*, *41*(3-4), 157-179.
- Hauser, M. D. (1991). Sources of Acoustic Variation in Rhesus Macaque (*Macaca mulatta*) Vocalizations. *Ethology*, *89*(1), 29-46.
- Hauser, M. D. (1998). Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Animal Behaviour*, *55*(6), 1647-1658.
- Hauser, M. D., and Marler, P. (1993). Food-associated calls in rhesus macaques (*Macaca mulatta*): I. Socioecological factors. *Behavioral Ecology*, *4*(3), 194-205.
- Hauser, M. D., Evans, C. S., and Marler, P. (1993). The role of articulation in the production of rhesus monkey, *Macaca mulatta*, vocalizations. *Animal behaviour*, *45*(3), 423-433.

- Hauser, M. D., and Ybarra, M. S. (1994). The role of lip configuration in monkey vocalizations: experiments using xylocaine as a nerve block. *Brain and language*, 46(2), 232-244.
- Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *science*, 298(5598), 1569-1579.
- Henzi, S. P. (1982). Some aspects of visual signalling and social organisation in the vervet monkey (*Cercopithecus Aethiops Pygerythrus*) (Doctoral dissertation, University of KwaZulu-Natal (South Africa)).
- Henzi, S. P., and Lucas, J. W. (1980). Observations on the inter-troop movement of adult vervet monkeys (*Cercopithecus aethiops*). *Folia primatologica*, 33(3), 220-235.
- Hinds, D. S., and Calder, W. A. (1971). Tracheal dead space in the respiration of birds. *Evolution*, 25(2), 429-440.
- Hohmann, G. (1991). Comparative analyses of age-and sex-specific patterns of vocal behaviour in four species of old world monkeys. *Folia Primatologica*, 56(3), 133-156.
- Holley, A. J. (1993). Do brown hares signal to foxes? *Ethology*, 94(1), 21-30.
- Hollien, H., Green, R., and Massey, K. (1994). Longitudinal research on adolescent voice change in males. *The Journal of the Acoustical Society of America*, 96(5), 2646-2654.
- Horrocks, J. (1986). Life-history characteristics of a wild population of vervets (*Cercopithecus aethiops sabaesus*) in Barbados, West Indies. *International Journal of Primatology*, 7(1), 31-47.
- Isbell, L. A., and Bidner, L. R. (2016). Vervet monkey (*Chlorocebus pygerythrus*) alarm calls to leopards (*Panthera pardus*) function as a predator deterrent. *Behaviour*, 153(5), 591-606.
- Isbell, L. A., Cheney, D. L., and Seyfarth, R. M. (2004). Why vervet monkeys (*Cercopithecus aethiops*) live in multimale groups. In R. H. Tuttle (Ed.), *The guenons: Diversity and adaptation in African monkeys* (pp. 173-187). Springer, Boston, MA.
- Jaccard, J., Becker, M. A., and Wood, G. (1984). Pairwise multiple comparison procedures: A review. *Psychological Bulletin*, 96(3), 589.
- Japkowicz, N. (2000). The class imbalance problem: Significance and strategies. In Proceedings of the 2000 International Conference on Artificial Intelligence (IC-AI'2000): Special Track on Inductive Learning Las Vegas, Nevada.
- Jarrett, J. D. (2019). *Vervet monkey development in a life historical perspective* (Doctoral dissertation, University of Lethbridge (Canada)).
- Junger, J., Pauly, K., Bröhr, S., Birkholz, P., Neuschaefer-Rube, C., Kohler, C. and Habel, U. (2013). Sex matters: Neural correlates of voice gender perception. *Neuroimage*, 79, 275-287.

- Jürgens, U. (1979). Vocalization as an emotional indicator A neuroethological study in the squirrel monkey. *Behaviour*, 69, 88-117.
- Jürgens, U. (1998). Neuronal control of mammalian vocalization, with special reference to the squirrel monkey. *Naturwissenschaften*, 85(8), 376-388.
- Jürgens, U., Hammerschmidt, K., and Fichtel, C. (2001). On the vocal expression of emotion. A multi-parametric analysis of different states of aversion in the squirrel monkey. *Behaviour*, 138(1), 97-116.
- Kappeler, P. M. (1990). Female dominance in *Lemur catta*: more than just female feeding priority? *Folia Primatologica*, 55(2), 92-95.
- Kappeler, P. M., and van Schaik, C. P. (2004). Sexual selection in primates: review and selective preview. In P. M. Kappler, C. P. van Schaik (Eds.) *Sexual selection in primates: new and comparative perspectives* (pp. 3-23), Cambridge University Press, Cambridge.
- Kassambara, A. (2017a). *Machine learning essentials: Practical guide in R* (Vol. 1). STHDA
- Kassambara, A. (2017b). *Practical guide to cluster analysis in R: Unsupervised machine learning* (Vol. 1). STHDA.
- Kaufman, L., and Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (Vol. 344). John Wiley and Sons
- Kelley, D. B. (2004). Vocal communication in frogs. *Current opinion in neurobiology*, 14(6), 751-757.
- Kent, J. P. (1987). Experiments on the relationship between the hen and chick (*Gallus gallus*): The role of the auditory mode in recognition and the effects of maternal separation. *Behaviour*, 102(1), 1-13.
- Kirchhof, J., and Hammerschmidt, K. (2006). Functionally referential alarm calls in tamarins (*Saguinus fuscicollis* and *Saguinus mystax*)—evidence from playback experiments. *Ethology*, 112(4), 346-354.
- Kiriazis, J., and Slobodchikoff, C. N. (2006). Perceptual specificity in the alarm calls of Gunnison's prairie dogs. *Behavioural Processes*, 73(1), 29-35.
- Kitchen, D. M. (2004). Alpha male black howler monkey responses to loud calls: effect of numeric odds, male companion behaviour and reproductive investment. *Animal Behaviour*, 67(1), 125-139.
- Kitchen, D. M., Seyfarth, R. M., Fischer, J., and Cheney, D. L. (2003). Loud calls as indicators of dominance in male baboons (*Papio cynocephalus ursinus*). *Behavioral ecology and sociobiology*, 53(6), 374-384.
- Kitzmann, C. D., and Caine, N. G. (2009). Marmoset (*Callithrix geoffroyi*) food-associated calls are functionally referential. *Ethology*, 115(5), 439-448.

- Kondo, N., and Watanabe, S. (2009). Contact calls: information and social function. *Japanese Psychological Research*, 51(3), 197-208.
- Krebs, J. R., and Dawkins, R. (1984). Animal signals: Mind reading and manipulation. In J. R. Krebs, and N. B. Davies (Eds.), *Behavioural ecology: An evolutionary approach* (2nd ed., pp.380–402). Oxford, England: Blackwell.
- Kruskal, W. H., and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260), 583-621.
- Kuhn (2018). Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team, Michael Benesty, Reynald Lescarbeau, Andrew Ziem, Luca Scrucca, Yuan Tang, Can Candan and Tyler Hunt. (2018). caret: Classification and Regression Training. R package version 6.0-81. <https://CRAN.R-project.org/package=caret>
- Lancaster, J. 1975. *Primate Behavior and the Emergence of Human Culture*. New York: Holt.
- Laukkanen, A. M., Vilkmann, E., Alku, P., and Oksanen, H. (1997). On the perception of emotions in speech: The role of voice quality. *Logopedics Phoniatrics Vocology*, 22(4), 157-168.
- Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22.
- Liebal, K., Waller, B. M., Slocombe, K. E., and Burrows, A. M. (2014). *Primate communication: a multimodal approach*. Cambridge University Press.
- Loughry, W. J., Oeser, M., & Hoogland, J. L. (2019). Alarm calls of the same individual vary during a response to the same predator in Gunnison's prairie dogs (*Cynomys gunnisoni*). *Canadian Journal of Zoology*, 97(11).
- Lubbe, A., Hetem, R. S., McFarland, R., Barrett, L., Henzi, P. S., Mitchell, D., and Fuller, A. (2014). Thermoregulatory plasticity in free-ranging vervet monkeys, *Chlorocebus pygerythrus*. *Journal of Comparative Physiology B*, 184(6), 799-809.
- Macedonia, J. M. (1990). What is communicated in the antipredator calls of lemurs: evidence from playback experiments with ringtailed and ruffed lemurs. *Ethology*, 86(3), 177-190.
- Macedonia, J. M., and Evans, C. S. (1993). Essay on contemporary issues in ethology: variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology*, 93(3), 177-197.
- MacLean, E. L., and Hare, B. (2012). Bonobos and chimpanzees infer the target of another's attention. *Animal Behaviour*, 83(2), 345-353.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K.(2018). cluster: Cluster Analysis Basics and Extensions. R package version 2.0.7-1.

- Maestriperi, D., and Call, J. (1996). Mother-infant communication in primates. In *Advances in the Study of Behavior* (Vol. 25, pp. 613-642). Academic Press.
- Maestriperi, D., Leoni, M., Raza, S. S., Hirsch, E. J., and Whitham, J. C. (2005). Female copulation calls in guinea baboons: Evidence for postcopulatory female choice?. *International Journal of Primatology*, 26(4), 737-758.
- Mann, H. B., and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60.
- Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1483), 2315-2324.
- Manser, M. B. (2010). The generation of functionally referential and motivational vocal signals in mammals. In *Handbook of behavioral neuroscience* (Vol. 19, pp. 477-486). Oxford. Elsevier.
- Manser, M. B., Bell, M. B., and Fletcher, L. B. (2001). The information that receivers extract from alarm calls in suricates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1484), 2485-2491.
- Manser, M. B., Seyfarth, R. M., and Cheney, D. L. (2002). Suricate alarm calls signal predator class and urgency. *Trends in cognitive sciences*, 6(2), 55-57.
- Marler, P. (1967). Animal Communication Signals: We are beginning to understand how the structure of animal signals relates to the function they serve. *Science*, 157(3790), 769-774.
- Marler, P., Evans, C. S., and Hauser, M. D. (1992). Animal signals: motivational, referential, or both. In Hanus Papousek, U Jurgens, and M. Papousek (Eds.), *Nonverbal vocal communication: Comparative and developmental approaches* (pp-66-86), Cambridge University Press.
- Martin, A., and Santos, L. R. (2016). What cognitive representations support primate theory of mind?. *Trends in cognitive sciences*, 20(5), 375-382.
- Masataka, N. (1987). The Perception of Sex-specificity in Long Calls of the Tamarin (*Saguinus labiatus labiatus*). *Ethology*, 76(1), 56-64.
- Masataka, N. (1994). Lack of correlation between body size and frequency of vocalizations in young female Japanese macaques (*Macaca fuscata*). *Folia Primatologica*. 63, 115-118.
- Masumoto, T. (1999). Size assortative mating and reproductive success of the funnel-web spider, *Agelena limbata* (Araneae; Agelenidae). *Journal of insect behavior*, 12(3), 353-361.
- McCowan, B., Franceschini, N. V., and Vicino, G. A. (2001). Age differences and developmental trends in alarm peep responses by squirrel monkeys (*Saimiri sciureus*). *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 53(1), 19-31.

- McElligott, A. G., Gammell, M. P., Harty, H. C., Paini, D. R., Murphy, D. T., Walsh, J. T., and Hayden, T. J. (2001). Sexual size dimorphism in fallow deer (*Dama dama*): do larger, heavier males gain greater mating success?. *Behavioral Ecology and Sociobiology*, 49(4), 266-272.
- McFarland, R., Fuller, A., Hetem, R. S., Mitchell, D., Maloney, S. K., Henzi, S. P., and Barrett, L. (2015). Social integration confers thermal benefits in a gregarious primate. *Journal of Animal Ecology*, 84(3), 871-878.
- Medvin MB, Stoddard PK, Beecher MD (1993) Signals for parent-offspring recognition: a comparative-analysis of the begging calls of the cliff swallows and barn swallows. *Animal Behaviour* 45(5), 841–850.
- Mehta, P. H., and Josephs, R. A. (2010). Testosterone and cortisol jointly regulate dominance: Evidence for a dual-hormone hypothesis. *Hormones and behavior*, 58(5), 898-906.
- Meise, K., Keller, C., Cowlshaw, G., and Fischer, J. (2011). Sources of acoustic variation: implications for production specificity and call categorization in chacma baboon (*Papio ursinus*) grunts. *The Journal of the Acoustical Society of America*, 129(3), 1631-1641.
- Melis, A. P., Call, J., and Tomasello, M. (2006). Chimpanzees (*Pan troglodytes*) conceal visual and auditory information from others. *Journal of Comparative Psychology*, 120(2), 154.
- Mercier, S., Neumann, C., van de Waal, E., Chollet, E., de Bellefon, J. M., and Zuberbühler, K. (2017). Vervet monkeys greet adult males during high-risk situations. *Animal Behaviour*, 132, 229-245.
- Miller, C. T., Scarl, J., and Hauser, M. D. (2004). Sensory biases underlie sex differences in tamarin long call structure. *Animal Behaviour*, 68(4), 713-720.
- Mitani, J. C. (1985). Sexual selection and adult male orangutan long calls. *Animal Behaviour*, 33(1), 272-283.
- Mitani, J. C., and Stuht, J. (1998). The evolution of nonhuman primate loud calls: acoustic adaptation for long-distance transmission. *Primates*, 39(2), 171-182.
- Moore, F., Mabey, S., and Woodrey, M. (2003). Priority access to food in migratory birds: age, sex and motivational asymmetries. In *Avian migration* (pp. 281-292). Springer, Berlin, Heidelberg.
- Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, 111(981), 855-869.
- Morton, E. S. (2017). *Animal vocal communication: assessment and management roles*. Cambridge University Press.
- Mullennix, J. W., Johnson, K. A., Topcu-Durgun, M., and Farnsworth, L. M. (1995). The perceptual representation of voice gender. *The Journal of the Acoustical Society of America*, 98(6), 3080-3095.

- Mundry, R., and Sommer, C. (2007). Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour*, 74(4), 965-976.
- Murray, I. R., and Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, 93(2), 1097-1108.
- Nesterova, N. L. (1996). Age-dependent alarm behavior and response to alarm call in bobac marmots (*Marmota bobac* Mull.). In M. Le Berre, R. Ramouse, and L. Le Guelte (Eds.), *Biodiversity in marmots* (pp. 181-186). International Network on Marmots, Moscow, Russia, and Lyon, France.
- Nuechterlein G. L., and Buitron, D. (1992). Vocal advertising and sex recognition in eared grebes. *The Condor*, 94(4), 937-943.
- O'Connell, S. M., and Cowlshaw, G. (1994). Infanticide avoidance, sperm competition and mate choice: the function of copulation calls in female baboons. *Animal behaviour*, 48(3), 687-694.
- Oliveira, D. A., and Ades, C. (2004). Long-distance calls in Neotropical primates. *Anais da Academia Brasileira de Ciências*, 76(2), 393-398.
- Owings, D. H., and Morton, E. S. (1998). *Animal vocal communication: a new approach*. Cambridge University Press.
- Owren, M. J., and Bernacki, R. H. (1988). The acoustic features of vervet monkey alarm calls. *The Journal of the Acoustical Society of America*, 83(5), 1927-1935.
- Owren, M. J., Berkowitz, M., and Bachorowski, J. A. (2007). Listeners judge talker sex more efficiently from male than from female vowels. *Perception and psychophysics*, 69(6), 930-941.
- Owren, M. J., and Rendall, D. (1997). An affect-conditioning model of nonhuman primate vocal signaling. In *Communication* (pp. 299-346). Springer, Boston, MA.
- Owren, M. J., and Rendall, D. (2001). Sound on the rebound: bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 10(2), 58-71.
- Owren, M. J., Rendall, D., and Bachorowski, J. A. (2005). Conscious and Unconscious Emotion in Nonlinguistic Vocal Communication.
- Owren, M. J., Seyfarth, R. M., and 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.
- Pasternak, G., Brown, L. R., Kienzle, S., Fuller, A., Barrett, L., and Henzi, S. P. (2013). Population ecology of vervet monkeys in a high latitude, semi-arid riparian woodland. *Koedoe*, 55(1), 01-09.

- Penn, D. C., and Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 731-744.
- Pereira, M. E., and Macedonia, J. M. (1991). Ringtailed lemur anti-predator calls denote predator class, not response urgency. *Animal Behaviour*, 41(3), 543-544.
- Pfefferle, D., and Fischer, J. (2006). Sounds and size: identification of acoustic variables that reflect body size in hamadryas baboons, *Papio hamadryas*. *Animal Behaviour*, 72(1), 43-51.
- Pfefferle, D., Heistermann, M., Hodges, J. K., and Fischer, J. (2008). Male Barbary macaques eavesdrop on mating outcome: a playback study. *Animal Behaviour*, 75(6), 1885-1891.
- Pitcher, B. J., Harcourt, R. G., and Charrier, I. (2010). Rapid onset of maternal vocal recognition in a colonially breeding mammal, the Australian sea lion. *PLoS One*, 5(8), e12195.
- Plotsky, K., Rendall, D., Riede, T., and Chase, K. (2013). Radiographic analysis of vocal tract length and its relation to overall body size in two canid species. *Journal of Zoology*, 291(1), 76-86.
- Pollermann, B. Z., and Archinard, M. (2002). Acoustic patterns of emotions. In E. Keller, G. Baily, A. Monaghan, J. Terken, and M. Huckvale (Eds.), *Improvements in speech synthesis* (pp. 237-245). Chichester: John Wiley & Sons.
- Poole, J. H. (1994). Sex differences in the behaviour of African elephants. In R. V. Short and E. Balaban (Eds.), *The differences between the sexes* (pp. 331-346). Cambridge, U.K.: Cambridge University Press.
- Pradhan, G. R., Engelhardt, A., van Schaik, C. P., and Maestripieri, D. (2006). The evolution of female copulation calls in primates: a review and a new model. *Behavioral Ecology and Sociobiology*, 59(3), 333-343.
- Premack, D. 1975. On the origins of language. In: M. S. Gazzaniga, C. B. Blakemore (Eds) *Handbook of Psychobiology* (pp. 591-605). New York: Academic Press.
- Price, T., and Fischer, J. (2014). Meaning attribution in the West African green monkey: influence of call type and context. *Animal cognition*, 17(2), 277-286.
- Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K., and Fischer, J. (2015). Vervets revisited: A quantitative analysis of alarm call structure and context specificity. *Scientific reports*, 5, 13220.
- Provine, R. R. (1992). Contagious laughter: Laughter is a sufficient stimulus for laughs and smiles. *Bulletin of the Psychonomic Society*, 30(1), 1-4.
- Puts, D. A., Doll, L. M., and Hill, A. K. (2014). Sexual selection on human voices. In *Evolutionary perspectives on human sexual psychology and behavior* (pp. 69-86). Springer, New York, NY.

- Puts, D. A., Gaulin, S. J., and Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and human behavior*, 27(4), 283-296.
- Puts, D. A., Hill, A. K., Bailey, D. H., Walker, R. S., Rendall, D., Wheatley, J. R., Welling, L.L., Dawood, K., Cárdenas, R., Burriss, R.P., Jablonski, N. G., Shriver, M. D., Weiss, D., Lameira, A. R., Apicella, C. L., Owren, M. J., Barelli, C., Glenn, M. E., and Ramos-Fernandez, G. (2016). Sexual selection on male vocal fundamental frequency in humans and other anthropoids. *Proceedings of the Royal Society B: Biological Sciences*, 283(1829), 20152830.
- Puts, D. A., Hodges, C. R., Cárdenas, R. A., and Gaulin, S. J. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340-344.
- Puts, D. A., Jones, B. C., and DeBruine, L. M. (2012). Sexual selection on human faces and voices. *Journal of sex research*, 49(2-3), 227-243.
- Ramakrishnan, U., and G. Coss, R. (2000). Age differences in the responses to adult and juvenile alarm calls by bonnet macaques (*Macaca radiata*). *Ethology*, 106(2), 131-144.
- Ramakrishnan, U., and Coss, R. G. (2001). Strategies used by bonnet macaques (*Macaca radiata*) to reduce predation risk while sleeping. *Primates*, 42(3), 193-206.
- Ramos-Fernández, G. (2005). Vocal communication in a fission-fusion society: do spider monkeys stay in touch with close associates? *International Journal of Primatology*, 26(5), 1077-1092.
- Randall, J. A., Matocq, M. D. (1997). Why do kangaroo rats (*Dipodomys spectabilis*) footdrum at snakes? *Behavioral Ecology*, 8(4), 404-413.
- Range, F., and Fischer, J. (2004). Vocal repertoire of sooty mangabeys (*Cercocebus torquatus atys*) in the Taï National Park. *Ethology*, 110(4), 301-321.
- 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.
- Reby, D., McComb, K., Cargnelutti, B., Darwin, C., Fitch, W. T., and 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.
- Rendall, D. (2003). Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *The Journal of the Acoustical Society of America*, 113(6), 3390-3402.
- Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (2005). Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: the role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America*, 117(2), 944-955.

- Rendall, D., Notman, H., and Owren, M. J. (2009b). Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. *The Journal of the Acoustical Society of America*, 125(3), 1792-1805.
- Rendall, D., Owren, M. J., and Rodman, P. S. (1998). The role of vocal tract filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalizations. *The Journal of the Acoustical Society of America*, 103(1), 602-614.
- Rendall, D., Owren, M. J., and Ryan, M. J. (2009a). What do animal signals mean? *Animal Behaviour*, 78(2), 233-240.
- Rendall, D., Owren, M. J., Weerts, E., and Hienz, R. D. (2004). Sex differences in the acoustic structure of vowel-like grunt vocalizations in baboons and their perceptual discrimination by baboon listeners. *The Journal of the Acoustical Society of America*, 115(1), 411-421.
- Rendall, D., Rodman, P. S., and Emond, R. E. (1996). Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Animal behaviour*, 51(5), 1007-1015.
- Rendall, D., Seyfarth, R. M., Cheney, D. L., and Owren, M. J. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour*, 57(3), 583-592.
- Riede, T., and Fitch, T. (1999). Vocal tract length and acoustics of vocalization in the domestic dog (*Canis familiaris*). *Journal of Experimental Biology*, 202(20), 2859-2867.
- Riede, T., Bronson, E., Hatzikirou, H., and Zuberbühler, K. (2005). Vocal production mechanisms in a non-human primate: morphological data and a model. *Journal of Human Evolution*, 48(1), 85-96.
- Robbins, R., and McCreery, E. K. (2003). African wild dog pup vocalizations with special reference to Morton's model. *Behaviour*, 140(3), 333-351.
- Robinson, S. R. (1981). Alarm communication in Belding's ground squirrels. *Zeitschrift für Tierpsychologie*, 56(2), 150-168.
- Rowell, T. E., and Hinde, R. A. (1962). Vocal communication by the rhesus monkey (*Macaca mulatta*). In *Proceedings of the Zoological Society of London* (Vol. 138, No. 2, pp. 279-294). Oxford, UK: Blackwell Publishing Ltd.
- Santos, L. R., Nissen, A. G., and Ferrugia, J. A. (2006). Rhesus monkeys, *Macaca mulatta*, know what others can and cannot hear. *Animal Behaviour*, 71(5), 1175-1181.
- Sanvito, S., Galimberti, F., and Miller, E. H. (2007). Vocal signalling of male southern elephant seals is honest but imprecise. *Animal Behaviour*, 73(2), 287-299.
- Sapolsky, R. M. (2017). *Behave: The biology of humans at our best and worst*. Penguin.
- Sayigh, L. S., Esch, H. C., Wells, R. S., and Janik, V. M. (2007). Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*. *Animal Behaviour*, 74(6), 1631-1642.

- Scarantino, A. (2010). Animal communication between information and influence. *Animal Behaviour*, 79(6), e1-e5.
- Scarantino, A. (2013). Rethinking functional reference. *Philosophy of Science*, 80(5), 1006-1018.
- Scarantino, A., and Clay, Z. (2015). Contextually variable signals can be functionally referential. *Animal Behaviour*, 100, e1-e8.
- Schel, A. M., Candiotti, A., and Zuberbühler, K. (2010). Predator-detering alarm call sequences in Guereza colobus monkeys are meaningful to conspecifics. *Animal Behaviour*, 80(5), 799-808.
- Schel, A. M., Tranquilli, S., and Zuberbühler, K. (2009). The alarm call system of two species of black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*). *Journal of Comparative Psychology*, 123(2), 136.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological bulletin*, 99(2), 143.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech communication*, 40(1-2), 227-256.
- Schmid, V. S., and de Vries, H. (2013). Finding a dominance order most consistent with a linear hierarchy: an improved algorithm for the IandSI method. *Animal Behaviour*, 86(5), 1097-1105.
- Schoen, M. A. (1971). The anatomy of the resonating mechanism in howling monkeys. *Folia Primatologica*, 15(1-2), 117-132.
- Schwartz, J. J. (1986). Male calling behavior and female choice in the neotropical treefrog *Hyla microcephala*. *Ethology*, 73(2), 116-127.
- Scott-Phillips, T. C. (2008). Defining biological communication. *Journal of evolutionary biology*, 21(2), 387-395.
- Searcy, W. A., and Nowicki, S. (2005). *The evolution of animal communication: reliability and deception in signaling systems*. Princeton University Press.
- Sebe, F., Nowak, R., Poindron, P., and Aubin, T. (2007). Establishment of vocal communication and discrimination between ewes and their lamb in the first two days after parturition. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology*, 49(4), 375-386.
- Semple, S. (1998). The function of Barbary macaque copulation calls. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1393), 287-291.
- Semple, S. (2001). Individuality and male discrimination of female copulation calls in the yellow baboon. *Animal behaviour*, 61(5), 1023-1028.

- Semple, S., and McComb, K. (2000). Perception of female reproductive state from vocal cues in a mammal species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1444), 707-712.
- Serrano-Meneses, M. A., Córdoba-Aguilar, A., Méndez, V., Layen, S. J., and Székely, T. (2007). Sexual size dimorphism in the American rubyspot: male body size predicts male competition and mating success. *Animal Behaviour*, 73(6), 987-997.
- Seyfarth, R. M., and Cheney, D. L. (1984). The acoustic features of vervet monkey grunts. *The Journal of the Acoustical Society of America*, 75(5), 1623-1628.
- Seyfarth, R. M., and Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, 34(6), 1640-1658.
- Seyfarth, R. M., and Cheney, D. L. (1990). The assessment by vervet monkeys of their own and another species' alarm calls. *Animal Behaviour*, 40(4), 754-764.
- Seyfarth, R. M., and Cheney, D. L. (1997). Some general features of vocal development in nonhuman primates. In C. Snowdon and Hausberger (Eds.) *Social influences on vocal development*, 249-273. Cambridge: Cambridge University Press
- Seyfarth, R. M., and Cheney, D. L. (2003). *Meaning and emotion in animal vocalizations*. In P. Ekman, J.J. Campos, R.J. Davidson, and F.B.M. de Waal (Eds.), *Emotions inside out: 130 years after Darwin's The Expression of the Emotions in Man and Animals*, *Annals of the New York Academy of Sciences* (Vol. 1000, pp. 32–55). New York: New York Academy of Sciences.
- Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980a). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, 210(4471), 801-803.
- Seyfarth, R. M., Cheney, D. L., and Marler, P. (1980b). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Animal Behaviour*, 28(4), 1070-1094.
- Seyfarth, R. M., Cheney, D. L., Bergman, T., Fischer, J., Zuberbühler, K., and Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Animal Behaviour*, 80(1), 3-8.
- Seyfarth, R. M., Cheney, D. L., Harcourt, A. H., and Stewart, K. J. (1994). The acoustic features of gorilla double grunts and their relation to behavior. *American Journal of Primatology*, 33(1), 31-50.
- Shelley, E. L., and Blumstein, D. T. (2004). The evolution of vocal alarm communication in rodents. *Behavioral Ecology*, 16(1), 169-177.
- Sieber, O. J. (1986). Acoustic recognition between mother and cubs in raccoons (*Procyon lotor*). *Behaviour*, 96(1-2), 130-162.

- Sieber, O. J. (1984). Vocal communication in raccoons (*Procyon lotor*). *Behaviour*, 90(1-3), 80-113.
- Silk, J. B., Kaldor, E., and Boyd, R. (2000). Cheap talk when interests conflict. *Animal Behaviour*, 59(2), 423-432.
- Silk, J. B., Rendall, D., Cheney, D. L., and Seyfarth, R. M. (2003). Natal attraction in adult female baboons (*Papio cynocephalus ursinus*) in the Moremi Reserve, Botswana. *Ethology*, 109(8), 627-644.
- Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistics Compass*, 3(2), 621-640.
- Slater, P. J., and Mann, N. I. (2004). Why do the females of many bird species sing in the tropics?. *Journal of Avian Biology*, 35(4), 289-294.
- Sloan, J. L., and Hare, J. F. (2008). The more the scarier: adult Richardson's ground squirrels (*Spermophilus richardsonii*) assess response urgency via the number of alarm signallers. *Ethology*, 114(5), 436-443.
- Slobodchikoff, C. N., and Placer, J. (2006). Acoustic structures in the alarm calls of Gunnison's prairie dogs. *The Journal of the Acoustical Society of America*, 119(5), 3153-3160.
- Slocombe, K. E., and Zuberbühler, K. (2005). Agonistic screams in wild chimpanzees (*Pan troglodytes schweinfurthii*) vary as a function of social role. *Journal of Comparative Psychology*, 119(1), 67.
- Slocombe, K. E., and Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology*, 15(19), 1779-1784.
- Slocombe, K. E., and Zuberbühler, K. (2006). Food-associated calls in chimpanzees: responses to food types or food preferences? *Animal behaviour*, 72(5), 989-999.
- Slocombe, K. E., and Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, 104(43), 17228-17233.
- Smith, A. S., Birnie, A. K., Lane, K. R., and French, J. A. (2009). Production and perception of sex differences in vocalizations of Wied's black-tufted-ear marmosets (*Callithrix kuhlii*). *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 71(4), 324-332
- Smith, J. M., and Harper, D. (2003). *Animal signals*. Oxford University Press.
- Smith, W.J. (1977). *The Behavior of Communicating*. Cambridge, Mass: Harvard University Press.

- Snowdon, C. T. (2004). Sexual Selection and Communication. In Kappeler, P., and van Schaik, C. P. (Eds.), *Sexual Selection in Primates: New and Comparative Perspectives*, (pp. 57-70), Cambridge University Press, Cambridge, UK.
- Sokhi, D. S., Hunter, M. D., Wilkinson, I. D., and Woodruff, P. W. (2005). Male and female voices activate distinct regions in the male brain. *Neuroimage*, 27(3), 572-578.
- Soltis, J., Leong, K., and Savage, A. (2005). African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. *Animal Behaviour*, 70(3), 589-599.
- Sproul, C., Palleroni, A., and Hauser, M. D. (2006). Cottontop tamarin, *Saguinus oedipus*, alarm calls contain sufficient information for recognition of individual identity. *Animal behaviour*, 72(6), 1379-1385.
- Sterck, E. H., Watts, D. P., and van Schaik, C. P. (1997). The evolution of female social relationships in nonhuman primates. *Behavioral ecology and sociobiology*, 41(5), 291-309.
- Struhsaker, T. T. (1967a). Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In S. A. Altmann (Ed.) *Social Communication among Primates* (pp. 281-324), Chicago, University of Chicago Press
- Struhsaker, T.T. (1967b). Social structure among vervet monkeys (*Cercopithecus aethiops*). *Behaviour*, 29, 83-121.
- Struhsaker, T.T. (1967c). Ecology of vervet monkeys (*Cercopithecus aethiops*) in the Masai - Amboseli Game Reserve, Kenya. *Ecology*, 48, 891-904.
- Swan, D. C., and Hare, J. F. (2008). Signaler and receiver ages do not affect responses to Richardson's ground squirrel alarm calls. *Journal of Mammalogy*, 89(4), 889-894.
- Taylor, A. M., and Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research. *Journal of Zoology*, 280(3), 221-236.
- Taylor, A. M., Reby, D., and McComb, K. (2008). Human listeners attend to size information in domestic dog growls. *The Journal of the Acoustical Society of America*, 123(5), 2903-2909.
- Temeles, E. J. (1994). The role of neighbours in territorial systems: when are they "dear enemies"? *Animal Behaviour*, 47(2), 339-350.
- Templeton, C. N., Greene, E., and Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science*, 308(5730), 1934-1937.
- Thompson, A. B., and Hare, J. F. (2010). Neighbourhood watch: multiple alarm callers communicate directional predator movement in Richardson's ground squirrels, *Spermophilus richardsonii*. *Animal Behaviour*, 80(2), 269-275.

- Tibbetts, E. A., and Dale, J. (2007). Individual recognition: it is good to be different. *Trends in ecology and evolution*, 22(10), 529-537.
- Tilson, R. L., and Norton, P. M. (1981). Alarm duetting and pursuit deterrence in an African antelope. *The American Naturalist*, 118(3), 455-462.
- Titze, I. R. (1994). Principles of Voice Production. Prentice-Hall. *Englewood Cliffs, NJ*.
- Tomasello, M., and Call, J. (1997). *Primate cognition*. Oxford University Press, USA.
- Townsend, S. W., and Manser, M. B. (2013). Functionally referential communication in mammals: the past, present and the future. *Ethology*, 119(1), 1-11.
- Townsend, S. W., Deschner, T., and Zuberbühler, K. (2008). Female chimpanzees use copulation calls flexibly to prevent social competition. *PLoS One*, 3(6), e2431.
- Turner, T. R., Anapol, F., and Jolly, C. J. (1997). Growth, development, and sexual dimorphism in vervet monkeys (*Cercopithecus aethiops*) at four sites in Kenya. *American Journal of Physical Anthropology*, 103(1), 19-35.
- Vannoni, E., and McElligott, A. G. (2007). Individual acoustic variation in fallow deer (*Dama dama*) common and harsh groans: a source-filter theory perspective. *Ethology*, 113(3), 223-234.
- Venables, W. N. and Ripley, B. D. (2002) *Modern Applied Statistics with S* (4th Edition) Springer, New York. ISBN 0-387-95457-0
- Waaramaa, T., Alku, P., and Laukkanen, A. M. (2006). The role of F3 in the vocal expression of emotions. *Logopedics Phoniatrics Vocology*, 31(4), 153-156.
- Waaramaa, T., Laukkanen, A. M., Airas, M., and Alku, P. (2010). Perception of emotional valences and activity levels from vowel segments of continuous speech. *Journal of voice*, 24(1), 30-38.
- Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., and Fischer, J. (2015). Characterizing vocal repertoires—Hard vs. soft classification approaches. *PloS one*, 10(4), e0125785.
- Warkentin, K. J., Keeley, A. T., and Hare, J. F. (2001). Repetitive calls of juvenile Richardson's ground squirrels (*Spermophilus richardsonii*) communicate response urgency. *Canadian Journal of Zoology*, 79(4), 569-573.
- Waser PM. 1982. The evolution of male loud calls among mangabeys and baboons. In: C. T. Snowdon, C. H. Brown, and M. R. Peterson (Eds.), *Primate communication* (pp. 117-144). Cambridge: Cambridge University Press.
- Wheeler, B. C. (2009). Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proceedings of the Royal Society B: Biological Sciences*, 276(1669), 3013-3018.

- Wheeler, B. C. (2010). Production and perception of situationally variable alarm calls in wild tufted capuchin monkeys (*Cebus apella nigritus*). *Behavioral ecology and sociobiology*, 64(6), 989-1000.
- Wheeler, B. C., and Fischer, J. (2012). Functionally referential signals: a promising paradigm whose time has passed. *Evolutionary Anthropology: Issues, News, and Reviews*, 21(5), 195-205.
- Wheeler, B. C., and Fischer, J. (2015). The blurred boundaries of functional reference: a response to Scarantino and Clay. *Animal Behaviour*, 100, e9-e13.
- Wich, S. A., and Nunn, C. L. (2002). Do male "long-distance calls" function in mate defense? A comparative study of long-distance calls in primates. *Behavioral Ecology and Sociobiology*, 52(6), 474-484.
- Willems, E. P., and Hill, R. A. (2009). Predator-specific landscapes of fear and resource distribution: effects on spatial range use. *Ecology*, 90(2), 546-555.
- Wilson, D. R., and Evans, C. S. (2008). Mating success increases alarm-calling effort in male fowl, *Gallus gallus*. *Animal Behaviour*, 76(6), 2029-2035.
- Wolfheim, J. H. (1983). *Primates of the world: distribution, abundance and conservation*. Psychology Press.
- Woodland, D. J., Jaafar, Z., and Knight, M. L. (1980). The "pursuit deterrent" function of alarm signals. *The American Naturalist*, 115(5), 748-753.
- Wrangham, R. W. (1981). Drinking competition in vervet monkeys. *Animal Behaviour*, 29(3), 904-910.
- Wrangham, R. W., and Waterman, P. G. (1981). Feeding behaviour of vervet monkeys on *Acacia tortilis* and *Acacia xanthophloea*: with special reference to reproductive strategies and tannin production. *The Journal of Animal Ecology*, 50(3), 715-731.
- Wyman, M. T., Mooring, M. S., McCowan, B., Penedo, M. C. T., and Hart, L. A. (2008). Amplitude of bison bellows reflects male quality, physical condition and motivation. *Animal Behaviour*, 76(5), 1625-1639.
- Wyman, M. T., Mooring, M. S., Mccowan, B., Penedo, M. C. T., Reby, D., and Hart, L. A. (2012). Acoustic cues to size and quality in the vocalizations of male North American bison, *Bison bison*. *Animal Behaviour*, 84(6), 1381-1391.
- Ybarra, M. A. S. (1986). Loud calls of adult male red howling monkeys (*Alouatta seniculus*). *Folia primatologica*, 47(4), 204-216.
- Ybarra, M. (1988). Morphological adaptations for loud phonation in the vocal organ of howling monkeys. *Primate Rep*, 22, 19-24.

- Yeomans, J. S., and Frankland, P. W. (1995). The acoustic startle reflex: neurons and connections. *Brain research reviews*, 21(3), 301-314.
- Young, C., McFarland, R., Barrett, L., and Henzi, S. P. (2017). Formidable females and the power trajectories of socially integrated male vervet monkeys. *Animal Behaviour*, 125, 61-67.
- Zollinger, S. A., and Brumm, H. (2015). Why birds sing loud songs and why they sometimes don't. *Animal Behaviour*, 105, 289-295.
- Zuberbühler, K. (2000). Referential labelling in Diana monkeys. *Animal Behaviour*, 59(5), 917-927.
- Zuberbühler, K. (2002). Effects of natural and sexual selection on the evolution of guenon loud calls. In R. H. Tuttle (Ed.), *The guenons: diversity and adaptation in African monkeys* (pp. 289-306). Springer, Boston, MA.
- Zuberbühler, K. (2003). Referential signaling in non-human primates: cognitive precursors and limitations for the evolution of language. In P. J. B. Slater, C. T. Snowdon, T. J. Roper, M. Naguib (Eds.), *Advances in the Study of Behavior* (vol 33, pp. 265-308). Oxford. Elsevier.
- Zuberbühler, K. (2006). Alarm calls – Evolutionary and cognitive mechanisms. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (2nd ed, pp. 143-155). Oxford: Elsevier.
- Zuberbühler, K., and Neumann, C. (2017). Referential communication in nonhuman animals. In /j. Call, G. M. Burghardt, I. M. Pepperberg, C. T. Snowdon, T. Zentall (Eds.). *APA handbook of comparative psychology: Basic concepts, methods, neural substrate, and behavior* (Vol. 1. pp. 645-661). American Psychological Association.
- Zuberbühler, K., Cheney, D. L., and Seyfarth, R. M. (1999). Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology*, 113(1), 33.
- Zuberbühler, K., and Fedurek, P. (2016). Vocal Grooming. In V. Weekes-Shackelford, T. Shackelford, V. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science*. Springer, Cham.
- Zuberbühler, K., Jenny, D., and Bshary, R. (1999). The predator deterrence function of primate alarm calls. *Ethology*, 105(6), 477-490.
- Zuberbühler, K., Noë, R., and Seyfarth, R. M. (1997). Diana monkey long-distance calls: messages for conspecifics and predators. *Animal Behaviour*, 53(3), 589-604.

Appendices

Appendix to Chapter 2

Table A.1: VERVET MONKEY ETHOGRAM

Samara game reserve, South Africa

(Colin Dubreuil 2015-2019)

SOLITARY BEHAVIOUR		
F	Feed/ Forage	Focal engages in any one of a variety of behaviours relating to the location, processing and consumption of food. Subject may be manipulating food with hands, placing food in mouth, digging in the dirt for food, or looking through a tree (or shrub) for food. Food searching behaviours may include digging, sweeping hand(s) side to side on ground while looking down (in search of food), and picking leaves/fruit from branches. Active stalking of insects both on a substrate or in the air is considered foraging. Active stalking is defined as following the movement of insects while approaching to eat, or directed movement towards an insect before catching/consuming them. Note: simply chewing food stored in cheek pouches is not considered foraging. Small (within 2 meters) bouts of travel within a tree/shrub while searching for food is still considered foraging. Record start and end time
FN	Forage near	Feeding/Foraging within 2 meters of another individual. Record ID of other individual. Record start and end times
K	Drink	Focal consumes water. Record start and end time
I	Inactive	Focal sits, lies down or hangs (eyes can be open or closed). Record start and end time
T	Travel	Focal moves (does not include moving around while foraging for food: limit of movement during foraging is 2m. Movement outside of this 2m limit during foraging is considered a travel). Record start and end time
TN	Travel near	Traveling within 2 meters of another individual. Travelers seem to be going in same general direction. Record ID of other individual. Record start and end times

AG	Auto groom	Focal grooms itself (does not include scratching). Animal uses lips, teeth, and/or fingers to remove items from its fur, hands, feet, anogenital region, teeth, nails, nose, ears, and eyes. Record start and end time
NSP	Non-social play	Solitary play, i.e. swinging from tail, playing with broken branch. Actions are repetitive, stereotyped, and do not seem to achieve another purpose. Record start and end time
VG	Vigilance	Subject is stationary, paying attention to its environment/some environmental/social occurrence (i.e., i.e. large raptor flies over head, animal detects another social group, animal is focused on the social interaction of other individuals in its social group, etc...). Alternatively, subject may be actively scanning its environment. Looks/scanning behaviour should last more than 4 seconds to be considered vigilance.
WC	Wound Clean	Focal cleans wound by touching and licking. Record start and end time
ITP	Inter troop encounter - Participate	Focal animal actively engaging in physical and vocal displays and threats towards members of another social group (or lone individual from outside their own social group). Record start and end time
SOCIAL BEHAVIOUR (indicate direct, receive, or mutual)		
N	Sit Near	Focal is within 2 meters (2 arms reach) of other individual(s). Record ID of other individual(s). Record start and end time
C	Sit Close	Focal is within 1 meter (but not in contact) with other individual(s). Indicate other individual(s). Record start and end time
B	Sit in Body Contact	Focal is in physical contact with other individual(s). Indicate other individual(s). Record start and end time
A	Approach	Focal directs (or receives) an approach within 2 meters of other individual (exclude approaches between mother and dependent offspring [=still carried by the mother]). Indicate other individual involved.
ADI	Approach/Displace	Focal directs (or receives) an approach towards another individual, "causing" the receiver to move away from the location he/she has been occupying. Indicate other individual involved.
L	Leave	Focal directs (or receives) a leave, i.e. one individual within 2m of another individual distance itself from that individual (moves

		outside the 2m radius). Exclude mother-dependent offspring dyads. Indicate other individual involved.
P	Pass By	Focal directs (or receives) a pass by, i.e. one individual passes within 2 meters of another and keeps on moving (again, leaving the 2m radius) without stopping. Indicate other individual involved.
G	Allogroom	Focal directs or receives allogroom from other individual (5 sec break between bouts). Animal uses lips, teeth, and/or fingers to remove items from the fur, hands, feet, anogenital region, nose, ears, and eyes of another individual. Indicate other individual involved and d/r. Record start and end time
SG	Solicit Groom	Focal directs (or receives) solicitation to allogroom, i.e. one individual presents body part to another individual for grooming. Record start and end time
P	Social Play	Focal plays with other individual (chasing, wrestling, mock biting and usually accompanied by play vocalization). Actions are repetitive, stereotyped, and do not seem to achieve another purpose. Indicate individuals involved. Record start and end time
N	Nurse	Focal nurses their offspring. Nursing is defined as when an infant suckles at its mother's nipple(s). Because it is difficult to tell if the infant is in actively "suckling", nursing includes all instances where the mother's nipple (s) are in the mouth of the infant. Record start and end time
ALC	Allocarry	An individual (not the mother) carries an infant/juvenile while moving. Record start and end time
IH	Infant Handle	An individual (not the mother) sits and handles an infant, letting infant climb on them and/or cling to. Record start and end time
SH	Solicit Infant Handle	Individual (not the mother) presents chest/stomach area to infant for it to climb on.
X	Copulation	Dorso-ventral mounting of a female by a male with pelvic thrusting. Male usually grips the back of the female's calf with his feet, and grips female's torso on the sides with hands. Indicate individuals involved. Indicate if ejaculator pause is observed, or if semen is clearly visible after mounting has ended. Record start and end time.
XR	Copulation resist	Female resists mating attempt of male. Male attempts to mount female (male may approach female from behind, sniff her

		anogenital region, place hands on her side in an attempt to mount), but female resists by moving away, or aggressing (see below) the male. Female is always directing this behaviour, while male is always the receiver
GI	Genital Inspect	Focal touches/sniffs anogenital region of another individual. Indicate other individual and d/r. Record start and end time
Z	Out of Sight	Focal is out of view. Record start and end time.
AGGRESSION		
AGG	Aggression	May or may not involve physical contact. Can include any of the behaviors listed below. Include start and end time.
EF	Eye flash	Eyebrows move quickly up and down, exposing light portion above eye. Directed towards another individual.
LU	Lunge	The monkey makes a sudden intense movement towards another monkey. It does not move over a large distance (<1m).
CH	Chase	The monkey chases another monkey partner at high speed. (Record start and end time). Distinguished from chases in social play, as aggressive chases are generally accompanied/preceded/followed by screams, barks, and other aggressive behaviours (see above and below). Unlike play chases, aggressive chases are no repetitive.
BI	Bite	Subject bites another monkey in the skin and fur - Subject seizes another monkey with the mouth/teeth/jaws
Additional contextual data		
AB	Alarm bout	Multiple monkeys are engaged in a bout of alarm calling. Does not have to include the focal animal. Record start and end time
ITE	Inter group encounter	Animals within the focal group see, and respond to animals of another social group. Record start and end time
GD	Game drive	Game drive vehicle drives by/stops to look at monkeys. Record start and end time
OPA	Open Area	Social group is moving from an area with abundant trees/shrubs into a relatively "open" area – An open area is defined as one where the average distance between the base of shrubs and trees exceeded 5m, and there was little to no canopy coverage overall.

Appendix to Chapter 4

Table A.2: Estimates (standardized beta coefficients) of the effects of increasing body weight on the Lowest Dominant Frequency Band (DF1) for grunts and alarms produced by males and females. Sigma “ σ ” values represent estimates of the effect of increasing body weight on the variability of DF1 for the calls of males and females.

Population-Level Effects:						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Intercept	1.74	0.08	1.56	1.88	933	1
σ Intercept	0.21	0.21	-0.18	0.62	483	1
Female Grunts	-2.25	0.08	-2.4	-2.08	940	1
Male Alarms	-0.39	0.09	-0.55	-0.2	983	1
Male Grunts	-2.23	0.08	-2.38	-2.05	928	1
Log Body weight	-1.31	0.42	-2.14	-0.5	723	1
Female Grunts: Log Body weight	1.03	0.42	0.22	1.86	719	1
Male Alarms: Log Body weight	-4.94	0.43	-5.76	-4.12	727	1
Male Grunts: Log Body weight	1.13	0.42	0.33	1.94	717	1
σ Female Grunts	-3.05	0.2	-3.43	-2.66	518	1
σ Male Alarms	-0.94	0.2	-1.35	-0.54	504	1.01
σ Male Grunts	-2.78	0.24	-3.22	-2.34	471	1
σ Log Body weight	-2.9	1.45	-5.73	-0.07	517	1.01
σ Female Grunts: Log Body weight	4.04	1.34	1.5	6.64	723	1
σ Male Alarms: Log Body weight	-3.27	1.56	-6.25	-0.08	537	1.01
σ Male Grunts: Log Body weight	4.82	1.51	1.93	7.9	577	1.01
Group-Level Effects:						
~Bout_recording (Number Of levels: 413)						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Intercept	0.02	0	0.01	0.02	346	1.01
σ Intercept	0.77	0.05	0.67	0.87	338	1.02
~Subject (Number Of levels: 105)						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Intercept	0.04	0	0.04	0.05	754	1
σ Intercept	0.65	0.07	0.54	0.8	546	1
Full model R ² values						
	Estimate	Est.Error	Q2.5	Q97.5		
R ² conditional	0.765	0.003	0.756	0.771		
R ² marginal	0.764	0.004	0.755	0.770		

Table A.3: Estimates (standardized beta coefficients) of the effects of increasing body weight on the Lowest Dominant Frequency Band (DFA2) for grunts and alarms produced by males and females. Sigma “ σ ” values represent estimates of the effect of increasing body weight on the variability of DFA2 for the calls of males and females.

Population-Level Effects:						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Intercept	1.23	0.08	1.07	1.38	999	1
σ Intercept	-0.51	0.12	-0.74	-0.26	1101	1
Female Grunts	-1.74	0.08	-1.89	-1.59	1103	1
Male Alarms	-0.05	0.08	-0.21	0.1	1333	1
Male Grunts	-1.82	0.09	-1.99	-1.64	822	1
Log Body weight	-0.37	0.41	-1.15	0.46	1336	1
Female Grunts: Log Body weight	0.25	0.42	-0.57	1.04	1701	1
Male Alarms: Log Body weight	-2.68	0.45	-3.54	-1.79	1341	1
Male Grunts: Log Body weight	0.09	0.43	-0.73	0.93	1381	1
σ Female Grunts	-0.29	0.13	-0.55	-0.05	1167	1
σ Male Alarms	-0.44	0.13	-0.71	-0.19	1514	1
σ Male Grunts	-0.4	0.14	-0.7	-0.13	1080	1
σ Log Body weight	-0.87	1.06	-2.95	1.15	1052	1
σ Female Grunts: Log Body weight	0.94	1.06	-1.07	3.07	1023	1
σ Male Alarms: Log Body weight	-1.19	1.13	-3.34	1.07	1033	1
σ Male Grunts: Log Body weight	0.97	1.12	-1.2	3.18	989	1
Group-Level Effects:						
~Bout_recording (Number Of levels: 413)						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Sd (Intercept)	0.22	0.02	0.18	0.26	689	1.01
Sd (σ Intercept)	0.38	0.04	0.3	0.46	643	1
~Subject (Number Of levels: 105)						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Sd (Intercept)	0.2	0.02	0.16	0.25	998	1
Sd (σ Intercept)	0.3	0.04	0.23	0.38	1098	1
Full model R ² values						
	Estimate	Est.Error	Q2.5	Q97.5		
R ² conditional	0.748	0.007	0.735	0.760		
R ² marginal	0.676	0.012	0.649	0.698		

Table A.4: Estimates (standardized beta coefficients) of the effects of increasing body weight on the call duration for grunts and alarms produced by males and females. Sigma “ σ ” values represent estimates of the effect of increasing body weight on the variability of call duration for the calls of males and females.

Population-Level Effects:						
	Estimate	Est.Error	l-95% CI	u-95% CI	Eff.Sample	\hat{R}
Intercept	-0.5	0.05	-0.59	-0.4	637	1
σ Intercept	-0.96	0.1	-1.17	-0.77	581	1.01
Female Grunts	0.22	0.05	0.12	0.32	707	1
Male Alarms	0.08	0.06	-0.03	0.19	656	1
Male Grunts	0.24	0.06	0.12	0.35	609	1
Log Body weight	0.26	0.33	-0.39	0.93	518	1
Female Grunts: Log Body weight	0.83	0.35	0.17	1.54	595	1
Male Alarms: Log Body weight	1.09	0.36	0.36	1.82	508	1
Male Grunts: Log Body weight	0.06	0.34	-0.62	0.73	504	1
σ Female Grunts	-0.11	0.11	-0.31	0.11	670	1.01
σ Male Alarms	-0.09	0.11	-0.29	0.13	735	1
σ Male Grunts	-0.04	0.12	-0.27	0.2	627	1.01
σ Log Body weight	2.59	0.9	0.89	4.34	321	1.01
σ Female Grunts: Log Body weight	-3.4	0.91	-5.21	-1.65	331	1.01
σ Male Alarms: Log Body weight	-0.55	0.96	-2.41	1.31	326	1.01
σ Male Grunts: Log Body weight	-3.25	0.93	-5.09	-1.42	268	1.01
Group-Level Effects:						
~Bout_recording (Number Of levels: 413)						
	Estimate	Est.Error	l-95% CI	u-95% CI	Eff.Sample	\hat{R}
Sd (Intercept)	0.12	0.02	0.09	0.16	412	1
Sd (σ Intercept)	0.3	0.03	0.24	0.36	767	1.01
~Subject (Number Of levels: 105)						
	Estimate	Est.Error	l-95% CI	u-95% CI	Eff.Sample	\hat{R}
Sd (Intercept)	0.11	0.02	0.08	0.14	534	1.01
Sd (σ Intercept)	0.18	0.04	0.1	0.26	415	1.01
Full model R ² values						
	Estimate	Est.Error	Q2.5	Q97.5		
R ² conditional	0.239	0.021	0.201	0.284		
R ² marginal	0.165	0.019	0.132	0.205		

Table A.5: Estimates (standardized beta coefficients) of the effects of increasing body weight on mean element length for grunts and alarms produced by males and females. Sigma “ σ ” values represent estimates of the effect of increasing body weight on the variability of the mean element length for the calls of males and females.

Population-Level Effects:						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Intercept	-0.5	0.05	-0.61	-0.4	645	1.01
σ Intercept	-1.32	0.14	-1.58	-1.04	749	1
Female Grunts	-0.05	0.05	-0.14	0.05	1168	1
Male Alarms	0.38	0.06	0.26	0.5	757	1.01
Male Grunts	0.06	0.06	-0.07	0.19	741	1.01
Log Body weight	0.17	0.33	-0.46	0.81	876	1
Female Grunts: Log Body weight	0.44	0.31	-0.17	1.03	1236	1
Male Alarms: Log Body weight	0.74	0.37	0.04	1.47	950	1
Male Grunts: Log Body weight	0.67	0.35	-0.04	1.35	840	1
σ Female Grunts	0.18	0.14	-0.11	0.44	802	1
σ Male Alarms	0.02	0.14	-0.27	0.3	972	1
σ Male Grunts	0.37	0.16	0.05	0.66	777	1
σ Log Body weight	1.56	1.12	-0.7	3.79	669	1
σ Female Grunts: Log Body weight	-1.35	1.13	-3.61	0.8	704	1
σ Male Alarms: Log Body weight	-3.06	1.19	-5.39	-0.61	657	1
σ Male Grunts: Log Body weight	-0.54	1.15	-2.82	1.76	669	1
Group-Level Effects:						
~Bout_recording (Number Of levels: 413)						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Sd (Intercept)	0.14	0.02	0.11	0.18	558	1
Sd (σ Intercept)	0.44	0.05	0.35	0.53	441	1
~Subject (Number Of levels: 105)						
	Estimate	Est.Error	1-95% CI	u-95% CI	Eff.Sample	\hat{R}
Sd (Intercept)	0.19	0.02	0.15	0.23	822	1
Sd (σ Intercept)	0.29	0.04	0.21	0.38	741	1.01
Full model R ² values						
	Estimate	Est.Error	Q2.5	Q97.5		
R ² conditional	0.348	0.049	0.284	0.489		
R ² marginal	0.207	0.019	0.172	0.245		