The Phylogeny and Function of Vocal Complexity in Geladas

by

Morgan L. Gustison

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
2017

Doctoral Committee:

Associate Professor Thore J. Bergman, Co-Chair
Associate Professor Jacinta C. Beehner, Co-Chair
Associate Professor Steven Abney
Associate Professor Elizabeth A. Tibbetts

May 10, 2016 | vol. 113 | no. 19 | pp. 5137–5458

PNAS

Proceedings of the National Academy of Sciences of the United States of America

www.pnas.org

Gelada vocal call patterns

Stem cell and placenta syncytiotrophoblast generation

Regulation of voltage-gated ion channels

Pelvic development in females

Anthropogenic disturbance and conservation

Morgan L. Gustison

gustison@umich.edu

ORCHID iD: 0000-0002-1162-8966

Dedication

This dissertation is dedicated to my late grandfather, Frank Altmaier. He would have gotten such a kick out of learning about geladas from the rooftop of Africa.

Acknowledgements

Learning how to be a scientist is not something a person can do alone. Several people contributed to my training these past six years, and I indebted to all of them. First and foremost, I want to thank Thore Bergman and Jacinta Beehner for bringing me into their academic family. Thore invested many hours mentoring me during weekly individual meetings, and both Thore and Jacinta invested many hours facilitating lab meetings, editing grant and manuscript drafts, writing reference letters, helping me to network, and providing field support. It was so wonderful to be an apprentice to two incredible scientific minds – I hope that some of their genius rubbed off on me. Additionally, I am so grateful that I could work at their field site in Ethiopia. Not only was stalking geladas was a ton a fun, but it was such a powerful experience to live in a culture that was so different than the one I grew up in. The many months I spent in Ethiopia changed my understanding of the world and my role in it. Going forward, I think I will be both a better scientist and a better philanthropist.

I am grateful to all other faculty who contributed to my learning at Michigan and abroad. Drs. Liz Tibbetts and Steven Abney were incredibly supportive committee members over the past few years who provided several ideas for the dissertation. Drs. Peter Fashing and Nga Nguyen direct the other gelada field site in Ethiopia, and I am grateful that they allowed me to come to their field site and collect acoustic recordings from their monkeys. Dr. Brandon Aragona took me in as an adjunct lab member when I was curious to know more about the neural mechanisms underpinning complex vocal communication. My experience in his lab was critical for preparing me for a neuroscience-focused post doc and in designing my larger research vision.

Additionally, I want to thank faculty who contributed to my learning prior to the PhD program. Without their mentorship, I would not even be here. Dr. Charles Snowdon was my undergraduate honors thesis advisor at UW-Madison. His advice to ‘follow your heart’ will stick with me forever, and it dictates nearly every decision I make. Chuck was the person who first encouraged me to try field work, which led to a thesis project with wild owl monkeys at Dr. Eduardo Fernandez-Duque’s field site in Argentina. Also, Dr. David Abbott (and his grad student Yves Aubert) directed fascinating research programs at the Primate Center in Madison. Working for these projects got me interested in primates and biomedical questions. Drs. Stuart Semple and Ann MacLarnon were amazing mentors during the MRes program at the University of Roehampton. The nuggets of wisdom I collected during my meetings with Stuart and the MRes courses enabled me to ‘hit the ground running’ at Michigan.

I want to thank all past and current University of Michigan Gelada Research Project (UMGRP) members for their hard work maintaining the research enterprise that is the Gelada Project. Specifically, I want to thank Marcela Benitez, Sofia Carrera, Megan Gomery, Lucy Ho, Julie Jarvey, Teera Losch, Amy Lu, David Pappano, Aliza le Roux, Eila Roberts, Sharmi Sen, Noah Snyder-Mackler for being such wonderful friends, mentors, science critics and field supporters. I also want to thank Shif Asrat, Ambaye Fanta, Setey Girmay, Esheti Jejaw, Tariku Waregay for continuously keeping the Simien Mountains field site up and running. Also, thank you to the army of undergraduates that helped process data: Chelsea Miller, Edna Chiang, Erin Person, Natalie Cook, Rebecca David, Jessica Feller, Niki Abdou, and Megan Skrzypek.

I am indebted to all Biopsychology (and Neuroscience) faculty, staff, and students for making the 4th floor of East Hall a fun and inspiring learning environment. The ‘pre-2011’ generations set the stage for my cohort, and I’m so glad to have many of them as friends and mentors. These amazing scientists include, but are not limited to: Katy Goldey, Jocelyn Richard, Eila Roberts, Ben Saunders, Shauna Resendez, Kirsten Porter-Stranksy, and Lindsay Yager. Also, as the JC Penney photoshoot proved, my cohort is the best: Daniel Castro, Liz Cogan, Jeff Pettibone, and Christian Peterson. I’m also thrilled to have many current Michigan students and staff as friends and colleagues. These amazing scientists include, but are not limited to: Crystal Carr, Sofia Carrera, Shannon Cole, Curtis Hazelwood, Jonte Jones, Ashley Kaiser, Alex Kawa, Erin Naffziger, Natalie Nevarez, Kyra Phillips, Pavlo Popov, Anne Sabol, Bryan Singer, Daria Tchessalova, Shelley Warlow, Sarah Westrick, and Katie Yoest.

Finally, I want to thank all my non-UM friends and family for their support throughout the program, despite knowing that I would become a doctor of philosophy and not medicine. Importantly, I want to thank Schuyler Cohen for being my best friend and friendliest critic. He always helps me see the bigger picture and not get caught up in the nitty-gritty details, and I love that we are on a never-ending quest to (a) find the perfect laugh and (b) make the world just a little bit better.

Table of Contents

List of Tables

List of Figures

Abstract

The complexity of vocal communication varies widely across taxa – from humans who can create an infinite repertoire of sound combinations to some non-human species that produce only a few discrete sounds. A growing body of research is aimed at understanding the origins of ‘vocal complexity’. And yet, we still understand little about the evolutionary processes that led to, and the selective advantages of engaging in, complex vocal behaviors. I contribute to this body of research by examining the phylogeny and function of vocal complexity in wild geladas (*Theropithecus gelada*), a primate known for its capacity to combine a suite of discrete sound types into varied sequences. First, I investigate the phylogeny of vocal complexity by comparing gelada vocal communication with that of their close baboon relatives and with humans. Comparisons of vocal repertoires reveal that geladas – specifically the males – produce a suite of unique or ‘derived’ call types that results in a more diversified vocal repertoire than baboons. Also, comparisons of acoustic properties reveal that geladas produce vocalizations with greater spectro-temporal modulation, a feature shared with human speech, than baboons. Additionally, I show that the same organizational principle – Menzerath’s law – underpins the structure of gelada vocal sequences (i.e., combinations of derived and homologous call types) and human sentences. Second, I investigate the function of vocal complexity by examining the perception of male complex vocal sequences (i.e., those with more derived call types), the contexts in which they are produced, and how their production differs across individuals. A playback experiment shows that female geladas perceive ‘complex’ and ‘simple’ vocal sequences as being different. Then, two observational studies show that male production of complex vocal sequences mediates their affiliative interactions with females, both during neutral periods and periods of uncertainty (e.g., following conflicts). Finally, I find evidence that vocal complexity can act as a signal of male ‘quality’, in that more dominant males exhibit higher levels of vocal complexity than their subordinate counterparts. Collectively, the work presented in this dissertation presents an integrative investigation of the ultimate origins of complex communication systems, and in the process, it highlights the critical importance of approaching the study of complexity from several scientific perspectives.

Chapter 1:

Introduction

‘Complexity’ is a feature that has evolved in a wide range of systems – from the organization of sub-atomic particles to the organization of entire ecosystems (S. E. Page, 2010; Wolfram, 2002). While difficult to directly quantify, complexity is thought to hover between randomness and predictability and characterizes systems with emergent properties that are greater than the sum of its parts (Kane & Higham, 2015; Lloyd & Pagels, 1988; S. E. Page, 2010). One of the most impressive complex systems is that of communication signals like spoken human language (Fitch, 2000, 2010; Shannon, 1948). An emergent property of language is that we can create an infinite array of meanings with an extensive, yet finite, repertoire of units like phonemes and words (Hauser, Chomsky, & Fitch, 2002; Pinker & Jackendoff, 2005). This is possible because of phonological, syntactical, semantic and pragmatic components (Fitch, 2010). In other words, humans can: (a) produce an extensive range of discrete sounds (phonology), (b) arrange these sounds into rule-governed sequences (syntax), (c) encode meaning into different combinations of sounds (semantics), and (d) modify these meanings based on context (pragmatics). There is an immense amount of interest in understanding the origins of complex vocal communication systems like spoken human language. Despite this great interest, however, why these systems have evolved remains a mystery (Hauser et al., 2014).

One solution to understanding the evolution of complex communication systems is to expand the comparative dataset from which we can draw inferences (Fitch, 2000; Hauser et al., 2014). Expanding our comparative dataset is also important because then we can better apply an ultimate perspective to the study of these complex vocal behaviors (Tinbergen, 1963). The ultimate perspective contains two components, phylogeny and function. “Phylogeny” refers to the evolutionary history of a behavior. The types of questions asked when exploring the phylogeny of a behavior include: (a) How has a behavior diverged between closely related species and converged amongst distantly related species? (b) What types of behaviors are shared or *homologous* across species and which are more recently evolved, or *derived*? “Function” refers to the adaptive value of a behavior. The types of questions asked when exploring the function of a behavior include: (a) What are the costs and benefits to using a behavior? and (b) What are the potential fitness consequences of a behavior?

In this dissertation, I describe a series of studies investigating the phylogeny and function of vocal complexity, using geladas (*Theropithecus gelada*) as a study model. Since captive and wild studies first began on them in the 1970s, geladas have been identified as a species with potentially unique vocal abilities (Aich, Moos-Heilen, & Zimmermann, 1990; Aiello & Dunbar, 1993; Dunbar & Dunbar, 1975; Richman, 1976, 1987). Specifically, geladas were shown to produce a large repertoire of sounds at incredibly high rates, and these sounds appeared to contain speech-like acoustic features (e.g., plosives, fricatives). Since the 1990s, little systematic work has followed up on these earlier observations. Meanwhile, there has been a wealth of literature published on the vocal abilities of the gelada’s closely related baboon relatives (e.g., (Rendall, Seyfarth, Cheney, & Owren, 1999)). In Chapters 2-4, I investigate the phylogeny of ‘complexity’ in gelada vocal communication by making comparisons between geladas and baboons, as well as between geladas and humans. Then, in Chapters 5-8, I investigate the functions of vocal complexity in gelada society.

Phylogeny of vocal complexity

Several comparative studies have investigated the components of vocal communication that have changed throughout recent evolutionary history in non-human taxa (Blumstein & Armitage, 1997; Krams, Krama, Freeberg, Kullberg, & Lucas, 2012; McComb & Semple, 2005; Pollard & Blumstein, 2012). One of the primary features of communication that has become more ‘complex’ in some species over others is vocal repertoire size, or the number of discrete sound types produced by a species. One of the explanations for this variation in vocal complexity comes from the ‘social complexity hypothesis’, which states that increases in some aspect of sociality – such as group size – promotes the need for a more diversified vocal communication system (Freeberg, 2006; Freeberg, Ord, & Dunbar, 2012). One of the prevailing questions that comes out of this comparative work is what parts of the vocal repertoire have diversified? This distinction matters because the answer results in different conclusions regarding how social complexity may co-evolve with vocal complexity. If, for example, the number of alarm calls are increased in species with large group sizes, then it would suggest that the diversified vocal repertoire is an outcome of greater predator threat that may come along with a living in a larger group. On the other hand, if the number of contact calls are increased in species with large group size, then it would suggest that more call types might function to help individuals coordinate their movements in the environment. To determine which components of vocal repertoires are diversified, it is important to compare closely related species. This maximizes the ability to figure out which vocalizations are homologous or conserved across taxa and which are unique or derived in a species. This line of thinking is what set up the rationale for Chapter 2.

In **Chapter 2** (Gustison, le Roux, & Bergman, 2012), I address the questions: (a) What components of primate vocal repertoires have diversified and (b) What are the implications for the social complexity hypothesis? I begin by making a qualitative exploration of the variation in vocal repertoires across a range of primate and non-primate taxa by separating vocal repertoires into six contextual based categories. The comparisons suggest that vocal repertoires have diversified in a modular way, with specific parts of the vocal repertoires diversified depending on the species. Second, I compare the vocal repertoire of geladas to baboons, a taxonomic group thought to have diverged from geladas only 4 mya (S. L. Page, Chiu, & Goodman, 1999). I compare these species because they are closely related and they differ greatly in standard socal complexity measures – geladas live in larger groups and spend more time socializing than do baboons. The gelada-baboon comparison shows that most of their vocal repertoires are shared. However, a suite of call types appears to be unique to gelada males. The male derived call types are relatively low-amplitude and tonal calls that appeared to fit in the same contextual category as ‘exhaled grunts’, a call type shared between geladas and other Papionins. These data imply that the gelada and baboon vocal repertoires differ because geladas evolved a diversified suite of ‘contact call’ vocalizations used in affiliative social interactions. I make some predictions about the functions of this subset of vocalizations and the resulting implications for the social complexity hypothesis. One of the limitations of this study, however, was that we were left with uncertainty about whether gelada vocalizations were also more acoustically complex than baboon vocalizations. Is it just the case that geladas make more types of vocalizations or are these vocalizations also more difficult to produce? This gap in knowledge is what set up the rationale for Chapter 3.

In **Chapter 3** (Gustison & Bergman, n.d.), I address the questions: (a) What spectral and temporal properties of gelada vocalizations have diverged from baboons? (b) Do the spectro-temporal properties of gelada vocalizations show any similarities with human speech? Currently, there are two main hypotheses on the vocal capabilities of non-human primates (Fitch, de Boer, Mathur, & Ghazanfar, 2016). The prevailing view – the "peripheral hypothesis" – states that non-human primates lack the anatomical ability to produce coordinated movements of the vocal tract that could produce modulated speech-like sounds. The alternative view – the "neural hypothesis" – states that non-human primates have the anatomical ability to produce complex sounds but lack the neural architecture needed to produce the sophisticated movements. In this chapter, I explore the capacity of geladas to modulate the spectrum based (e.g., fundamental frequency and formant bandwidths) and temporal based (i.e., cyclic patterns in wobbles) properties of their calls. First, I compare the exhaled grunts of geladas and baboons. Then, I compare gelada male exhaled grunts to the 'derived' call types described in Chapter 2. Overall, I found that gelada grunts show higher levels of vocal modulation than baboons, and this degree of modulation was even more extreme in some of the derived call types (exhaled moans, wobbles and yawns). I make the argument that the data provide further support for the neural hypothesis.

Chapters 2 and 3 focus on describing ‘vocal complexity’ in terms of repertoire size. Another feature of communication that seems to have become more ‘complex’ in some species over others is the ability to combine sounds together into variable sequences (Kershenbaum et al., 2014). Generally, animals rarely combine more than a few sounds together and for those that do, these sequences tend to be stereotyped. This means that the same sound or same pattern is repeated across vocal utterances. It is actually quite rare for animals to produce non-stereotyped sound combinations (for some exceptions see Kershenbaum, Ilany, Blaustein, & Geffen, 2012; Okanoya, 2004). Upon revisiting the older literature on gelada vocal behavior (Richman, 1987) and our database of acoustic recordings, I realized that male geladas were not producing exhaled grunts and derived call types on separate occasions, rather, they were combining these different sounds into varied sequences ranging anywhere from a single call to over 20 calls. By comparison, baboons rarely combine more than six grunts together, and they do not combine multiple call types (Rendall et al., 1999). These observations led me to ask whether similar organizing principles predict the structure of gelada and human communication systems, and this is the rationale for Chapter 4.

In **Chapter 4** (Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016), I test the hypothesis that organizing principles of human language – specifically ‘Menzerath’s Law’ – underpin the structure of gelada call sequences**.** Menzerath’s law predicts that constituent size (e.g., call duration) is negatively associated with construct size (e.g., sequence size in number of calls). I found evidence for this organizing principle in gelada male call sequences in that longer sequences tended to be composed of shortened calls (e.g., exhaled and inhaled grunts) and shorter sequences tended to contain prolonged calls (e.g., exhaled moans, wobbles and yawns). Taken together, these findings suggest that specific organizing principles may dictate how taxa evolved complex forms of combinatorial vocal systems. In addition, the data show that when describing 'vocal complexity' in geladas, we must look at the level of vocal sequences and not just individual vocalizations.

Function of vocal complexity

The findings of Chapters 2-4 are helpful for understanding the divergent and convergent evolutionary processes underlying gelada vocal complexity, but we still lack an understanding of the adaptive value of vocal complexity. In other words, how does vocal complexity benefit male geladas? I seek to answer this overarching question in Chapters 5-8. My first step in tackling this broad question was to better understand the receiver's point of view. Communication systems are '2-way' in that they involve the production of a message from a sender and the perception of that message from a receiver (Bradbury & Vehrencamp, 1998; Shannon, 1948). Although it is easy for human scientists to categorize vocal sequences as 'complex' or not, this does not guarantee that this categorization has ecological validity. To other geladas, derived call types may be perceived as redundant with exhaled grunts. On the other hand, derived call types may modify the 'meaning' of vocal sequences in some way. If the makeup of vocal sequences matters, then we would expect that receivers respond differently to sequences based on their composition. This line of thinking is what set up the rationale for Chapter 5.

In **Chapter 5** (Gustison & Bergman, 2016), I test the hypothesis that female geladas differentially respond to gelada male vocal sequences based on whether or not they contain one of the three most acoustically ‘elaborate’ derived call types – exhaled moans, wobbles and vocal yawns (hereto called the “derived call types”). We focused on females because they appear to be the intended receivers for these vocal sequences. This is because gelada social systems are harem-like and sequences are produced almost exclusively by males. To test whether females respond differently to sequences based on their composition, I used an experimental playback design in which I played male vocal sequences to females and monitored their acute responses to the sounds. Specifically, I looked at orientation behavior (i.e., looking time) and proximity behavior (i.e., time spent close to the speaker). Females looked longer at and spent more time in proximity to playbacks of sequences that contained one of the derived call types. The findings support the hypothesis that females perceive derived call vocal sequences differently from grunt-only sequences. These findings also open the door to a suite of follow up questions regarding the *contexts* in which complex sequences are produced. In Chapter 2, I proposed that the evolution of derived call types in geladas may have something to do with ‘social complexity’. I follow up on this hypothesis in Chapter 6.

In **Chapter 6**, I test the hypothesis that vocal complexity has specific functions related to measures of social complexity in geladas – specifically large group size and high levels of gregariousness. Challenges faced by species living in large groups include high levels of conspecific noise and difficulty maintaining visual/vocal contact with social partners. Highly gregarious species are faced with the challenge of needing effective strategies to form and maintain social bonds. In this chapter, I explore whether sequence production and complexity are predicted by features of the social environment (e.g., level of conspecific noise) or types of social interactions (e.g., approaches and grooming). I also investigate whether sequences are directed at females and whether producing a sequence leads to affiliative interactions with females. The data support the hypothesis that social complexity may lead to vocal complexity in geladas. Specifically, more complex sequences were produced by males during noisy conditions and during a wide range of affiliative interactions. Vocal sequences were directed at females and the use of more derived call types resulted in female investment in the social bond via grooming behavior. Taken together, these findings largely support the predictions made in chapter 2 (Gustison et al., 2012), that the need to maintain bonds within a noisy backdrop of conspecific vocalizations may favor greater vocal complexity. One of the limitations of this chapter is that the behavioral observations typically occurred during relatively stable periods in the male-female relationships. If vocal complexity functions to facilitate and maintain social relationships, then this vocal strategy should be particularly important when the stability of these relationships is at stake. This remaining question was the rationale for carrying out the study in Chapter 7.

In **Chapter 7**, I test the hypothesis that vocal complexity is the strategy that male geladas prefer to engage in following conflicts with other members of their society. There is a wealth of evidence to suggest that non-human animals engage in tactile-based strategies (e.g., grooming) during reconciliation and consolation interactions following conflicts (Aureli, 2002; McFarland & Majolo, 2013). Yet, as outlined in the Chapter, there are several reasons why vocal strategies may be preferred. For example, individuals that need to maintain social bonds with several partners (e.g., gelada males) would benefit from strategies that reach several individuals at once rather than one at a time (like in grooming). To test this hypothesis, I quantified the non-vocal and vocal behaviors that occurred following conflicts. Then, I used a pseudo-experimental design to track male vocal behavior following conflicts and during control periods. I found that males preferentially engage in vocal over tactile post-conflict strategies, while females were the opposite. I also found that male vocal sequences following conflicts were longer and more complex following conflicts than during control periods. Moreover, more complex sequences were more predictive of subsequent female investment in the relationship (female to male grooming).

Taken together, the findings of Chapters 5, 6 and 7 support the hypothesis that gelada male vocal complexity is particularly important in maintaining social bonds with females. The conclusions of these studies also fit more broadly in the discourse on sexual selection of extravagant traits (Johnstone, 1995; Zahavi, 1975). The findings of these chapters provide good evidence to suggest that gelada male vocal complexity evolved through inter-sexual selection. For an extravagant trait to be selected for in this capacity, however, it must be an honest advertisement of a male's quality. Without this piece of the puzzle, it is difficult to conclude that inter-sexual selection could promote an extravagant behavior such as gelada vocal complexity. This missing puzzle piece was the rationale for Chapter 8.

In **Chapter 8**, I test the hypothesis that vocal complexity mirrors male quality in geladas. First, I compare sequence production and complexity in dominant (leader) males and subordinate (follower) males. Second, I investigate changes in sequence production and complexity throughout leader male tenure, with the assumption that new leaders are higher quality than older leaders. Third, I investigate whether sequence production and complexity in the beginning of a leader's tenure predicts direct and indirect measures of fitness (e.g., units size, tenure length and number of births). I found that higher quality males produce more complex sequences than lower quality males, but no evidence that there were clear associations between vocal complexity and fitness measures. Overall, the findings support the hypothesis that vocal complexity is an honest signal of male quality. Combined with chapters 5-7, there is evidence to suggest that inter-sexual selection drove the evolution of vocal complexity in geladas.

Taken together, the studies in this dissertation reveal several answers regarding the phylogeny and function of vocal complexity in geladas. In **Chapter 9**, I summarize the implications of these answers and discuss remaining lines of inquiry to be tackled in future research.

References

Aich, H., Moos-Heilen, R., & Zimmermann, E. (1990). Vocalizations of adult gelada baboons (*Theropithecus gelada*): Acoustic structure and behavioural context. *Folia Primatologica*, *55*, 109–132.

Aiello, L. C., & Dunbar, R. I. M. (1993). Neocortex size, group size, and the evolution of language. *Current Anthropology*, *34*, 184–193.

Aureli, F. (2002). Conflict resolution following aggression in gregarious animals: a predictive framework. *Animal Behaviour*, *64*, 325–343.

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

Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of Animal Communication*. Sunderland, MA: Sinauer.

Dunbar, R. I. M., & Dunbar, P. (1975). *Social Dynamics of Gelada Baboons*. Basel, Switzerland: Karger.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, *4*, 258–267.

Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, *2*, e1600723.

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

Freeberg, T. M., Ord, T. J., & Dunbar, R. I. M. (2012). The social network and communicative complexity: preface to theme issue. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1782–1784.

Gustison, M. L., & Bergman, T. J. (In Revision). Divergent acoustic properties of gelada and baboon vocalizations and their implications for the evolution of human speech. *Journal of Language Evolution*.

Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Scientific Reports*, *6*, 19680.

Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (Theropithecus gelada) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1847–59.

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*, E2750–E2758.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., … Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, *5*, 1–12.

Johnstone, R. A. (1995). Honest advertisement of multiple qualities using multiple signals. *Journal of Theoretical Biology*, *177*, 87–94.

Kane, E. A., & Higham, T. E. (2015). Complex systems are more than the sum of their parts: Using integration to understand performance, biomechanics, and diversity. *Integrative and Comparative Biology*, *55*, 146–165.

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

Kershenbaum, A., Ilany, A., Blaustein, L., & Geffen, E. (2012). Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 2974–81.

Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: a parid perspective. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1879–91.

Lloyd, S., & Pagels, H. (1988). Complexity as thermodynamic depth. *Annals of Physics*, *188*, 186–213.

McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, *1*, 381–5.

McFarland, R., & Majolo, B. (2013). The importance of considering the behavioral form of reconciliation in studies of conflict resolution. *International Journal of Primatology*, *34*, 15–29.

Okanoya, K. (2004). The Bengalese finch: A window on the behavioral neurobiology of birdsong syntax. *Annals of the New York Academy of Sciences*, *1016*, 724–735.

Page, S. E. (2010). *Diversity and Complexity*. Princeton, New Jersey: Princeton University Press.

Page, S. L., Chiu, C., & Goodman, M. (1999). Molecular phylogeny of Old World monkeys (Cercopithecidae) as inferred from gamma-globin DNA sequences. *Molecular Phylogenetics and Evolution*, *13*, 348–59.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition*, *95*, 201–36.

Pollard, K. a, & Blumstein, D. T. (2012). Evolving communicative complexity: insights from rodents and beyond. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1869–78.

Rendall, D., Seyfarth, R. M., Cheney, D. L., & Owren, M. J. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour*, *57*, 583–592.

Richman, B. (1976). Some vocal distinctive features used by gelada monkeys. *The Journal of the Acoustical Society of America*, *60*, 718–24.

Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates*, *28*, 199–223.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System's Technical Journal*, *27*, 379–423.

Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift Fur Tierpsychology*, *20*, 410–433.

Wolfram, S. (2002). *A New Kind of Science*. Champaign, IL: Wolfram Media.

Zahavi, A. (1975). Mate selection-A selection for a handicap. *Journal of Theoretical Biology*, *53*, 205–214.

Chapter 2:

Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates

Abstract

Primates are intensely social and exhibit extreme variation in social structure, making them particularly well suited for uncovering evolutionary connections between sociality and vocal complexity. Although comparative studies find a correlation between social and vocal complexity, the function of large vocal repertoires in more complex societies remains unclear. We compared the vocal complexity found in primates to both mammals in general and human language in particular and found that non-human primates are not unusual in the complexity of their vocal repertoires. To better understand the function of vocal complexity within primates, we compared two closely-related primates (chacma baboons and geladas) that differ in their ecology and social structures. A key difference is that gelada males form long-term bonds with the 2-12 females in their harem-like reproductive unit, while chacma males primarily form temporary consortships with females. We identified homologous and non-homologous calls and related the use of the derived non-homologous calls to specific social situations. We found that the socially-complex (but ecologically-simple) geladas have larger vocal repertoires. Derived vocalizations of geladas were primarily used by leader males in affiliative interactions with 'their' females. The derived calls were frequently

used following fights within the unit suggesting that maintaining cross-sex bonds within a reproductive unit contributed to this instance of evolved vocal complexity. Thus, our comparison highlights the utility of using closely-related species to better understand the function of vocal complexity.

Introduction

The complexity of vocal communication varies enormously across species, from humans with an endless repertoire of sound combinations, to species of mongoose that produce only three different sounds [1]. As we continue to document the diversity that exists in nature, we are increasingly able to use comparative studies to identify the selective pressures responsible for increasing vocal complexity. One of the most salient findings that has emerged is that high levels of sociality are found in combination with a high degree of vocal complexity (*e.g.* [2–5]). For example, ground-dwelling sciurid species with socially complex groups (*e.g.,* number of age/sex classes) produce more acoustically distinct alarm calls than species with fewer age/sex classes [2]. Although non-primate taxa can be excellent study subjects for investigating the evolution of vocal complexity in general (e.g., rodents: [6]; bats [5]; primates, as our closest relatives, can provide insight into the evolution of the most complex vocal system—our own [7,8]). Moreover, primates exhibit extreme variation in social structure, making them particularly well suited for uncovering evolutionary connections between sociality and vocal complexity.

Facets of primates' sociality distinguish them from most other mammals. First, primates exhibit an unusual degree of sociality that some have proposed has resulted in a kind of ‘Machiavellian intelligence’ [9,10] in that individuals are capable of forming coalitions [11], deceiving others [12], and maintaining strong, long-term social bonds with both kin and non-kin [13–15]. Second, primates are unusual among mammals in that the size of their groups is positively associated with some aspects of brain size [10,13]. One intriguing explanation for this relationship is that primates require highly sophisticated cognitive abilities for keeping track of and maintaining complex networks of social relationships [9,16,17] – particularly considering recent data that indicate such social networks actually enhance individuals’ fitness [18,19]. Offering further support, the strong positive relationship between group size, time spent grooming and diversity of vocal repertoire in primates [20] suggests that more vocalisations may indeed be necessary for navigating the complex network of social relationships in primate societies.

In this review, we first focus on the evolution of vocal complexity in primates, and then propose a novel approach for studying the function of vocal complexity. Although a species’ repertoire size provides one useful comparative metric, it is a composite measure with no information about the function of the individual calls that comprise it. Here, we propose that the function of vocal complexity can be understood by comparing vocalisations among closely-related species with differing repertoire sizes to identify species-specific derived calls. In such cases, the function of greater vocal complexity equates to the function of the derived calls. As an example of how this approach can provide insights about the evolution of complexity, we compare the vocalisations of geladas (*Theropithecus gelada*) and chacma baboons (*Papio ursinus*) – two closely-related Old World monkeys with overlapping vocal repertoires but very different ecological and social structures.

Vocal repertoires of primates and other mammals

(a) Repertoire size as a measure of complexity

Mammalian vocal communication is typically described as being made up of discrete, functional units, or “calls” [21–23]. Based on these functional units, vocal complexity is quantified in terms of (a) number of discrete vocalisations in the repertoire (repertoire size) (*e.g.*, [2,20,24,25], or, less commonly, (b) degree of individuality within discrete calls [5,6,26]). Other ways of assessing vocal complexity include quantifying syllable complexity, amount of information contained within a call [5], or the number of calls within a specific category of vocalizations (alarm calls: [2]).

Several mammalian species produce call variants, or *graded* calls, which vary slightly in acoustic properties [27–31], such as fundamental frequency [28], ‘pitch’ [29] or duration [24] and, as a result, have different meanings to receivers [29,32–34]. For species with small, fixed vocal repertoires, these subtle alterations may help to extend the flexibility of an otherwise limited repertoire [7]. However, identifying graded calls requires detailed acoustic and behavioural analyses and data of comparable detail are rarely available for multiple species. Therefore, as has become the convention in vocal studies [2,20,24,25], we refer to repertoire size as the number of *discrete* calls that animals in a population or species produce.

(b) Vocalisation types

Vocalisations are produced in many different contexts. Some are produced in response to external stimuli such as predators and food. We call these ‘allospecific’ vocalizations (Table 2.1) and include alarm calls and food calls. Alarm calls can communicate the degree of risk involved [2,24], indicate predator type (*i.e.,* aerial or terrestrial) [35,36], or combine information on risk and predator type (*e.g.*, [37]). Primates, in particular, are known to produce alarm vocalizations specific to predator-type, eliciting appropriate responses in receivers [38–40]. Notably, the complexity of primate alarm calls is generally attributed to a complex physical (rather than social) environment [41]. If different predators have different modes of hunting, primate prey should have evolved different predator responses to each. By contrast, the complexity of alarm calls in a small social carnivore, the meerkat (*Suricata suricatta*), has been attributed to the need for social coordination [4]. Relative to a sympatric-living herbivore species like Cape ground squirrels (*Xerus inauris*), meerkats travel farther from underground shelters in their open habitat to find living prey, and they depend on ‘sentinels’ to emit referential alarm calls that vary acoustically based on predator type [4]. This strategy allows individual meerkats to decrease time spent being vigilant and increase foraging efficiency.

The other allospecific vocalisations, food calls, are less variable than alarm calls. Only a handful of studies have demonstrated that variation in calls are related to the quantity or quality of the food source (*e.g.,* [42]); but most studies report a *lack* of variation (*e.g.,* [43]). Although alarm and food calls differ substantially in degree of complexity, they share two features: (1) they are elicited by non-conspecifics, and (2) they are the only two contexts where proto-syntax (i.e., the combination of call elements to form new meanings) has been reported, specifically in primates (food calls [44]; alarm calls [45]).

The vast majority of mammalian vocalizations are emitted during social interactions with conspecifics, under conditions of varying motivational states (*e.g.,* mating, aggression, fear). We call these ‘social’ vocalisations and divide them into two main classes – calls that function over a long distance (‘loud calls’) and calls produced in close-range social interactions (‘close-range calls’, Table 2.1). Loud calls may function to attract or defend mating partners [46], defend a territory or food source through maintenance of intergroup spacing [47,48], or re-establish contact with group members that are out of sight (‘separation calls’) [49].

Close-range calls are produced in agonistic, neutral, or affiliative contexts. Calls produced in agonistic contexts may function to assess or warn rivals, such as contest calls that advertise fitness [48] or threat calls that maintain a dominance hierarchy [50,51]. Harassed individuals, on the other hand, may produce distress calls, which likely function to appease the aggressor and attract coalition partners [52]. The specific function of some close-range calls made in strictly neutral or affiliative social situations has been more difficult to ascertain and little is known about them besides the contexts in which they are produced [7]. Copulation calls, for example, do not appear to have the same function across different species and may even serve no function in some instances [53,54]. Other close-range calls are ascribed an affiliative function, often described as ‘contact calls’ (*e.g.*, in raccoons (*Procyon lotor*) [51]; capuchins (*Cebus capucinus*) [55]). Contact calls can be produced in various ‘friendly’ contexts, such as during post-conflict reconciliation interactions [56] and prior to friendly behaviour like allogrooming [52,57,58]. These close-range contact calls are also produced in more ‘neutral’ behavioural states like foraging and resting and therefore could be involved in the maintenance of group cohesion and inter-individual spacing [59].

Primate vocal repertoires are similar to those of other terrestrial mammals (Table 2.1). Although primate repertoires may be slightly larger (on average), there is considerable overlap between primates and other taxa, both in total repertoire size and within each category of calls. Species with large repertoires relative to other species in their Order generally produce a large proportion of calls in just one or two categories of calls (*e.g.,* long distance and competitive calls – *Callicebus moloch* [49]; distress and contact calls – *Pan troglodytes* [50]; allospecific and contact calls – *Suricata suricatta* [37,60]). This suggests that specific needs related to one domain (e.g., competition or affiliation) might drive the development of large repertoires, rather than an overall increase in repertoire across all categories. Within primates, no clear taxonomic pattern has emerged with respect to repertoire size. Each family of primates (including great apes) contains species with large and small repertoires. Surprisingly, despite the social complexity of primates, there is no consistent trend for primates to have more social calls than other mammals, which suggests that simple comparisons of numbers of calls are of limited utility.

(c) Function of larger repertoires

One of the primary hypotheses put forward to explain large, complex vocal repertoires is that social complexity creates the need for more vocalisations [21,61–64]. Comparative studies have found a positive relationship between social complexity and communicative complexity, providing support for this hypothesis [2,3,5,20]. In sciurids, the alarm call repertoire size increases with the number of demographic 'roles' [2]. Additionally, in primates, an increase in total vocal repertoire size was associated with both larger groups and increases in time spent grooming – a measure of social cohesion [20]. These studies have been important for pinpointing aspects of sociality (*i.e.,* large sociable groups, various demographic roles) that may drive the evolution of large repertoires. However, vocal repertoires of different species may be 'large' for different reasons (Table 2.1), and more work is clearly needed to understand the selective pressures underlying expansions in repertoires.

Vocal complexity of humans and other primates

Relative to humans, nonhuman primates (henceforth referred to as 'primates') exhibit surprisingly simplistic vocal production [65,66]. (Note that a focus on vocal *production* ignores the more sophisticated language-like abilities that primates exhibit in terms of vocal *perception* [67]). According to the 'source-filter-theory' first developed to describe human speech [68,69], vocal production entails two components: the 'source' of a vocalisation (*i.e.,* lungs and the vocal folds) and the means by which a vocalisation is shaped, or 'filtered', in the vocal cavities (*i.e.,* vocal tract). Speech relies heavily on the control of 'formants' or vocal resonances (a product of vocal tract morphology) to produce distinct syllables and hence encode information [69]. Primates also produce formants but the formant structure (*i.e.,* distance between sound frequency 'peaks and valleys' [69]) mainly encodes limited information such as individual identity [70,71] and body size [72–74]. Even more elaborate are humans' filtering tools, the descended larynx and tongue [69]. In most primates and other mammals, the tongue remains flat inside of the mouth. By contrast, humans have remarkable control over the location and shape of the tongue [69,75], giving humans unmatched plasticity in sound invention [76]. This unique vocal plasticity allows us to imitate complex sounds and invent novel sounds, a feat shared with some birds and cetaceans [77,78] but not with other primates [65,66].

Despite having a limited ability to imitate and create new sounds, there are features of primate vocal production that show similarities to human language. For instance, some primates exhibit vocal dialects – geographic variation in the acoustic structure of certain vocalizations [79–81]. Calls are recognizably homologous between different populations of the same species, but show acoustic distinctions related to variation in habitat and the duration of isolation, similar to patterns in human linguistic diversity (*e.g.,* [82]). Additionally, primates such as chimpanzees (*Pan troglodytes*) [83,84], blue monkeys (*Cercopithecus mitis*) [85] and capuchins (*Cebus apella*) [86,87] produce or suppress vocalizations depending on the composition of the conspecific audience.

Primate communication also resembles the semantic content of human language. Several primates exhibit potentially 'referential' allospecific calls that are elicited by external stimuli (Table 2.1). In some cases, the referential nature of these calls has been supported with playback experiments. For example, each vervet monkey (*Chlorocebus pygerythrus*) alarm call 'refers' to a different type of predator (leopards, eagles, and snakes). Experimental playbacks in the wild indicate that these different alarm calls produce different predator-appropriate responses in the absence of a predator [39]. In further support of the functionally-referential nature of primate vocalizations, habituation-dishabituation experiments on Diana monkeys (*Cercopithecus aethiops*) demonstrated that playbacks of leopard alarm calls *or* leopard growls resulted in predator-appropriate responses. These results suggest that Diana monkey responses are based on the underlying referent (the predator) rather than any differences in the calls' acoustic properties [88].

Despite some language-like properties of primate communication, humans exhibit unrivaled flexibility in mixing and matching different sounds to create new meanings through syntax [65,66,89]. Very few mammalian species use combinations of calls, and even those that do are unlikely to use these combinations to generate new meanings. There are only a few rare cases where primates were found to combine calls in ways that change the meaning of the call elements (red-capped mangabeys (*Cercocebus torquatus*) [90]; Campbell's monkeys (*Cercopithecus campbelli*) [44]; Diana monkeys (*Cercopithecus diana*) [40]). Importantly, these semantic combinations of sounds only comprise a few specific elements and are highly constrained [90,91]. Non-human primate vocal 'productivity' [92] is therefore far simpler than human communication and may, at best, be labelled as 'proto-syntax' – a term that refers to rule-governed, rather than random, combinations of discrete sounds that lack the sophistication of human grammar [44].

Function of derived vocalisations

Although previous studies have been pivotal for identifying aspects of sociality that drive vocal complexity, we still know relatively little about *how* large vocal repertoires function in complex societies. One reason for this is that comparisons of repertoire size alone fail to identify the specific calls that may have evolved in association with social complexity. With no knowledge on which calls are derived, we can say nothing about how those calls function. Another reason is that comparisons of group size alone fail to identify the specific features about group life that require an increase in vocal complexity. Thus, several questions remain unanswered: First, what specific aspects of sociality create a need for vocal complexity? Is it the number of relationships, the nature of relationships, or something else? Second, can we identify the derived components of the vocal repertoire that relate to the demands of increased sociality? That is, if more social species have more calls, how are they using these 'extra' calls?

To help answer these questions, we propose a systematic investigation of closely related species that make detailed comparisons of the functions of 'homologous' (shared between species) and 'derived' (unique to a species) vocalisations. Note that, although a vocalization may be unique to a species because it was present in the common ancestor and lost in the other species, we call them derived calls for simplicity, although the direction of the change (gain or loss) remains a hypothesis that can be examined by comparison to an outgroup. Previous studies have used comparisons among closely-related species to understand vocal evolution (e.g., [58]) although not with the goal of understanding vocal complexity *per se*. In the primate literature, several researchers have made comparisons of vocal behaviour between related species (*e.g.*, [93–96]). These studies often include general similarities and differences of call categories [93], acoustic structure [94], and/or contextual use [95]. In one case [95], there was a clear attempt to identify homologous and unique calls in two species of macaque (bonnet macaque (*Macaca radiata*) and lion-tailed macaque (*M. silenus*)) and two species of langurs (Nilgiri langur, *Presbytis johnii*) and common langur (*P. entellus*)), however much of the ensuing analyses focused on the differential use of the homologous calls rather than explaining the function of unique calls. The only study to date to focus on unique calls [97], compared the vocal behavior of the forest-dwelling mandrill (*Mandrillus sphinx*) to published accounts of savannah-living baboons (*Papio* spp.) and geladas (*Theropithecus gelada*). Kudo [97] reported that mandrills produced two unique long distance contact calls instead of the various short-range calls made by baboons and geladas. He proposed that this difference was likely due to ecological pressures, as low-amplitude vocalisations do not travel well in a forested environment where visual contact is also limited [97].

Identifying homologous and derived vocalisations is critical for identifying the specific social or ecological factors that may account for complex vocal repertoires. Here, we use a comparison of the vocal complexity in geladas with that in chacma baboons to demonstrate how this homologous-derived vocalisation strategy may be implemented. By analysing calls from both species (all obtained from wild populations under natural conditions), we control for variability in how calls are classified which may drive some of the variation in overall repertoire size found in meta-analyses.

Geladas and baboons – A case study

Early researchers were struck by the intricate vocal behaviour of geladas as well as their unusually-complex social groups [98–100]. Although some have proposed a causal connection between these factors [98], little progress has been made towards understanding why geladas, the only extant *Theropithecus* species, have elaborate vocal communication compared to other primates. Thus, a comparison between the vocal behaviour of geladas and *Papio* baboons serves two purposes. First, these two taxa split relatively recently (about 4 million years ago), and *Theropithecus* and *Papio* are likely sister genera [101]. To human observers, they appear to produce similar calls in similar contexts (e.g. affiliative grunts, threat grunts, and alarm calls). It is therefore relatively straightforward to identify homologous calls, and simultaneously, to pinpoint the unique call types that result in differences in vocal repertoire size. We can then assess how these unique calls are used, to highlight the selective pressures that may have favoured greater vocal complexity.

Second, the differences in the social system and ecology of geladas and baboons make the comparison particularly useful for testing contrasting predictions about the evolution of behavioural differences [102]. Both species live in matrilineal groups in which males disperse [99,103,104]. Geladas aggregate into a multi-level, fission-fusion society (forming groups as large as 1100 individuals) [105,106] and within this group they only recognize and primarily interact with a small subset of the individuals within ‘harem-like’ reproductive units of 2-15 individuals [99,105–108]. In these reproductive units, ‘leader’ males must maintain social relationships with several females, and it is thought that maintaining close social bonds with his unit females may serve to decrease the likelihood that he will be out-competed by a non-unit, ‘bachelor’ male [109,110]. In contrast to their complex social system, gelada diets are simple and specialized, with grass as the primary food item [111–113].

Unlike geladas, many baboons (*e.g.,* chacma baboons (*Papio ursinus*)) have a single-level, multimale-multifemale society with no discrete reproductive units (20-120 animals, [114–116]). Baboons maintain differentiated relationships based on kinship and dominance with all members of their group, but cross-sex relationships consist mainly of temporary consortships [114–117]. In terms of ecology, baboons are extremely complex; they live in a range of habitat types and consume anything from fruits and seeds to insects and vertebrates [118–121].

Given that geladas and baboons differ in their sociality and ecology, we predict corresponding differences in the call types comprising their vocal repertoires. For instance, geladas – specifically males – may produce more types of calls that are used in affiliative situations. On the other hand, baboons may use proportionally more allospecific calls to communicate about general features of the environment such as food items. To test our predictions, we compared the vocal behaviour of geladas with one representative of the *Papio* genus – the chacma baboon [97] – to identify derived call types. While we recognize that vocalizations from a single population may obscure variation within the genus, both the literature and our experience with multiple types of *Papio* baboons suggests that repertoire variation within *Papio* is minimal [97] and that the types of vocalizations used by chacma baboons are very similar to even the socially-divergent *Papio* species, *P. hamadryas* [97,122]. Furthermore, our descriptions of gelada vocalisations closely match those from captive geladas [123] suggesting that such vocalisations extend beyond those unique to one population. For any derived vocalisations, we then conducted intra-specific analyses to determine their possible functions.

(a) Study subjects

Data for this study come from 14 units within 3 different bands in one community of wild geladas (about 1200 individuals) living in the Sankaber area of the Simien Mountains National Park, Ethiopia (2008-2010) [104,105] and a single group of chacma baboons (group C) living in the Moremi Game Reserve in the Okavango Delta of Botswana (2001-2002). The gelada units were comprised of one leader male, 0-3 follower males, and 1-11 females and their immature offspring. The gelada habitat consisted of high-elevation open grassland and adjacent escarpments (sleeping sites). The chacma baboon group ranged from 82-91 individuals, including 9-11 adult males, 29-31 adult females, and their immature offspring. The baboon habitat was patchy scrub forest interspersed with seasonally-flooded grasslands.

(b) Comparison of gelada and chacma vocal repertoires

<u>Data analysis</u>

We opportunistically recorded vocalisations from 81 adult geladas (males=36; Feb 2008 – Apr 2010) and 32 adult chacma baboons (males=11; Apr 2001 – May 2002) with a Sennheiser ME66 directional microphone connected to a digital stereo recorder (Marantz PMD 660 Digital Recorder for geladas; Sony VW-D6 Professional Walkman for chacma baboons). The call types and contexts of all vocalisations were described at the time of recording. Our analyses focus on common calls that occurred repeatedly during focal sampling and we do not attempt to describe all vocalisations produced in each species. The inter-observer reliability (between assignments made in the field and assignments that were blind to previous designations and assignments based on isolated calls in the absence of contextual information) of a subset of these calls (5 exemplars/call type/species/sex class) was 96.0%. We used Avisoft (version 5.1.12, R. Specht, Berlin) to generate spectrograms with a fast fourier transformation size of 1024 points. Focusing on spectrograms with high signal-to-noise ratio, we categorised call types by ear, visual inspection of the spectrograms and the contexts in which they occurred (chacma females=50 calls; chacma males=32 calls; gelada females=72 calls; gelada males=92 calls). There were an equal number of calls per individual (within species/sex class) for each call type (n=1-3 call replicates per individual, 6-12 total calls per call type). We optimised the frequency range of different call types (11 or 22 kHz) where appropriate (time resolution of 2.667-2.903 ms and a 100% frame).

We used Avisoft to quantify 8 temporal and spectrum-based acoustic parameters in the spectrograms: duration, mean bandwidth, frequency under which 25% of the call's energy lies (start, maximum and mean), number of harmonic peaks under 20 dB (maximum and mean), maximum peak frequency. Then, to determine the probability of correctly assigning each vocalisation to a pre-categorized call type, we performed stepwise discriminant function analyses (DFAs) with a subsequent leave-on-out cross-validation procedure for each of the 4 species/sex classes separately [124]. We used multivariate analyses of variance (MANOVAs) to verify the significance of the final DFA parameters. Finally, we identified homologous calls between species based on both acoustic and contextual similarity.

Results

Male and female geladas and chacma baboons produced a range of allospecific and social calls used in both affiliative contexts (e.g., grooming and copulation) and non-affiliative contexts (e.g., challenge displays and dominance interactions), with geladas producing a greater number of call types (Table 2.2). Of the 14 call types we identified, 8 were found in both geladas and chacmas and 6 were unique to geladas. The derived gelada calls occurred primarily in short-range affiliative contexts (Table 2.2). Extant literature and our own observations indicated that most of the homologous call types are produced in a similar morphological way – a vocalised exhale – while geladas produce both inhaled and exhaled versions of calls that are acoustically distinct (we only separate inhaled and exhaled grunts here because they are the most common, but they also produce inhaled 'moans' and 'wobbles', Table 2.2).

We performed further analyses on 12 vocalisation types, of which only 7 were found in chacmas (Figures 2.1-2.2). Other call types were excluded from further DFA analyses because they were rarely produced without overlapping vocalisations, and hence, there were too few high-quality recordings to analyse (gelada female: display calls, moans, inhaled grunts, wobbles and yawns; gelada male: how barks, nasal inhaled grunts and alarm calls; chacma female: alarm calls; chacma male: fear grunt, alarm calls and copulation calls). We were able to discriminate between all call types for each age-sex class, using DFAs; based on 8 acoustic parameters, we classified call types at a higher rate (range: 67.4-93.8%; leave-one-out classification range: 50-90.6%) than expected by random classification (range: 10-33.3%). A MANOVA test carried out for each of the 4 species/sex classes showed that pre-categorized call types were significantly different from each other based on variation in at least 4 of the chosen acoustic parameters ($p < 0.003$).

In sum, acoustic analysis shows geladas share a number of vocalisation types with chacma baboons. While chacma baboons did not appear to have any unique calls, the analysis allowed identification of at least 5 derived vocalisation types in geladas: inhaled grunts, moans, pre-copulation calls, wobble calls and yawns. We then carried out intraspecific analysis to determine how these calls function in gelada society.

(c) Intraspecific analysis of derived gelada vocalisations

Comparison of derived call use in gelada males and females

To determine the function of the derived gelada vocalisations identified above, we first examined potential sex differences in the use of these calls. By definition, pre-copulation calls were produced only by females in very straightforward contexts (*i.e.,* produced prior to copulation). Thus, we focused here on the use of inhaled grunts, moans, wobbles and yawns for which the use of these calls is less obvious. Behavioural data on adult male and female geladas were obtained between January 2009 and December 2010 during repeated 15-minute focal follows of 53 females (mean ± SD: 6.55 h ± 2.59 h per female; 348.50 h in total) and 13 leader males (6.60 ± 1.86 h per male, 85.75 h in total). During these focal follows we noted all vocalisations uttered by the focal animal, as well as all social behaviour (*e.g.,* approaches and grooming interactions) involving the focal individual.

Next, we determined sex differences in the use of derived vocalisations by carrying out a general linear model (GLM) with sex and average reproductive unit size (over the entire study period) as fixed factors. We found that gelada males produced four of the derived calls (*i.e.,* inhaled grunts, moans, wobbles and yawns) at a higher mean rate (14.13 calls/h) than did gelada females (0.39 calls/h) ($F_{(1,63)}$=708.144, p < 0.001), and reproductive unit size did not come out as a significant covariate ($F_{(1,63)}$=0.942, p=0.336). Thus, males appeared to be the sex utilising derived vocalisations, and as such, we next explored whether these unique calls were used in contexts that are unique to males in gelada society.

Functionality of derived gelada male calls

First, we tested the hypothesis that derived social calls are used by males to maintain social relationships within the unit females by examining vocal behaviour in the context of conflict resolution. Using all adult female focal data, we identified every fight (both as actor and receiver) in which the focal animal was involved. These fights (N=107 events) were characterized by loud screams from the focal animal (n=48 events), or direct, physical attacks from the focal animal that included biting and slapping (n=59 events). We deliberately excluded any inconspicuous agonistic interactions (such as soft threat calls or visual threats) that may have gone unnoticed by other group members. For each fight event we counted all derived vocalizations directed at the focal animal by males in the 2 minutes preceding the event and the 2 minutes following the event and compared these values with binomial tests of proportions.

Second, we tested the hypothesis that derived social calls were used by males in association with the presence of non-unit, ‘bachelor’ males that pose a threat to the leader males (all leader males are eventually ousted by bachelor males). We used all adult leader male focal data for which the location of bachelors was stable throughout the entire 15-minute focal sample. In other words, bachelor groups were either close to the focal male (within 20 m: N=16 focals), far (more than 20 m; N=24 focals), or out of sight (N=26 focals). We carried out two GLMs with male identity as a random factor, bachelor distance as a fixed factor, and the rate of derived calls as the dependent variable (first model: close vs. far; second model: close vs. out of sight).

We found evidence that males used non-homologous derived calls to maintain cross-sex social relationships with females in his units. Specifically, we found that males directed derived non-homologous calls at females after fights happened (14 occurrences), and they never used them before a fight (binomial test of proportions: X^2_1= 12.916, $p < 0.001$). On the other hand, we did not find any evidence that males used the derived calls in response to the presence of bachelors. Leader males did not produce derived calls at a high rate when bachelor groups were close (3.23 calls/h) compared to when they were far away (2.04 calls/h) ($F_{(1,8)}=0.394$, $p=0.548$). Similarly, leader males did not produce derived calls at a higher rate when bachelor groups were close compared to when they were out of sight (2.51 calls/h) ($F_{(1,11)}=0.078$, $p=0.785$).

(d) General discussion

Geladas have an elaborate, almost ‘choral’ vocal repertoire [100] and live in a complex society with social groups of varying sizes making geladas an important model system for addressing hypotheses about vocal evolution. Identifying homologous vocalisations shared with *Papio* baboons allowed us to study the function of geladas derived vocalisations. It did not appear that interacting with many individuals [20,125] was necessarily an important factor in the use of derived calls, as their production was not correlated with the size of the reproductive unit. Rather, the need to maintain long-term bonds within the unit seemed most important; leader males used these derived vocalisations after fights broke out within their units. Thus, the gelada-specific vocalisations may have evolved as an adaptations to simultaneously maintaining relationships both with and among multiple females – leader males that are better able to ‘maintain the peace’ of their reproductive units may, in turn, have better reproductive success [109]. It remains to be determined why the cross-sex bonds seen in geladas seem more tightly linked to vocal complexity than the within-sex bonds found in both species.

Our results suggest that future studies should examine whether hamadryas baboons (*Papio anubis*), a *Papio* species that also has a ‘harem-like’ structure [126], have any evidence of greater elaboration of affiliative call use by males. This comparison is particularly important for uncovering how vocalizations relate to specific aspects of long-term bonds because hamadryas males form long-term bonds with females but the relationship is more coercive than in geladas and there does not appear to be a need to ‘keep the peace’. In geladas, investigations of how females respond to derived vocalisations and the subsequent benefits to leader male fitness is an exciting direction for future research. It may be the case, for instance, that these derived vocalisations directly benefit fitness by reducing female anxiety, similar to the proposed anxiolytic effects of grooming [127–129].

One puzzling aspect of our findings is that the derived calls used by males are all used in similar contexts. Further work is needed to tease apart any potential differences between the derived gelada calls but this redundancy suggests an additional hypothesis. Perhaps the extremely large groups of geladas (herds can number of 1,000 individuals) and high rates of vocalizations (mean ± SD: chacmas: 8.84 ± 4.49 calls/h, geladas: 16.95 ± 8.51 calls/h) create ‘vocal clutter’ that the geladas have overcome by diversifying their most common call types—affiliative vocalizations. Thus the need to maintain bonds within a noisy backdrop of conspecific vocalisations may favour greater vocal complexity, possibly explaining some of the group size effects seen in other studies [20].

Conclusions

Comparisons of repertoire size and components found that primates are broadly similar to other mammals, despite primates having greater social complexity. However, our comparison of baboons and geladas highlights the utility of making detailed comparisons among closely related species to understand vocal evolution. We were able to examine the function of recently-evolved calls in detail and examine the specific social implications of increased repertoires by focusing on specific call types, addressing sexual differences, and using behavioural measures to describe social complexity. We found that the larger repertoire of geladas is linked to the maintenance of cross-sex bonds within the reproductive unit. Broadly focused theoretical and comparative analyses [2,3,5,20] are vital to drive the investigation of communicative complexity, but we would argue that there is also a need for more focused analyses among carefully chosen taxa using directly comparable measures in the study of vocal complexity.

Acknowledgements

We are grateful to the Ethiopian Wildlife Conservation Department, the Amhara National Regional State Parks Development and Protection Authority, and the wardens and staff of the Simien Mountains National Park for granting us the permission to conduct this research in Ethiopia. We thank the Office of the President and the Department of Wildlife and National Parks of the Republic of Botswana for permission to conduct research in the Moremi Reserve. We would like to thank Esheti Jejaw, David Pappano, Eila Roberts, Noah Snyder-Mackler, and Vanessa Wilson for their assistance with collecting behavioural data in the field. Tiffany Fritzler and Chelsea Miller helped with the acoustic analysis and Ken Guire provided expert statistical advice. Research in Botswana was supported by NIH grant MH62249, an NRSA fellowship, the Leakey Foundation, and the University of Pennsylvania. Research in Ethiopia was supported by the Wildlife Conservation Society [SSF Gr# 67250], the National Geographic Society [Gr# 8100-06], the Leakey Foundation, the National Science Foundation [Gr# BCS-0715179], and the University of Michigan. We thank Dr Robin Dunbar, Dr Jacinta Beehner and two anonymous reviewers for helpful comments on a previous version of this manuscript.

References

1 Baker, C. 1988 Vocalization of captive water mongooses, Atilax paludinosus. *Z Saugetierkd* **53**, 83-91.

2 Blumstein, D. T. & Armitage, K. B. 1997 Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *Am Nat* **150**, 179-200. (DOI 10.1086/286062)

3 Freeberg, T. M. 2006 Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychol Sci* **17**, 557-561. (DOI 10.1111/j.1467-9280.2006.01743.x)

4 Furrer, R. D. & Manser, M. B. 2009 The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *Am Nat* **173**, 400-410. (DOI 10.1086/596541)

5 Wilkinson, G. S. 2003 Social and vocal complexity in bats. In *Animal social complexity: Intelli- gence, culture, and individualized societies* (eds F. B. M. de Waal & P. L. Tyack), pp. 322–341. Cambridge, MA: Harvard University Press.

6 Pollard, K. A. & Blumstein, D. T. this issue *Phil Trans R Soc B*

7 Fedurek, P. & Slocombe, K. E. 2011 Primate vocal communication: a useful tool for understanding human speech and language evolution? *Hum Biol* **83**, 153-173. (DOI 10.3378/027.083.0202)

8 Pinker, S. & Bloom, P. 1990 Natural-language and natural selection. *Behav Brain Sci* **13**, 707-726. (DOI 10.1017/S0140525X00081061)

9 Byrne, R. W. & Whiten, A. 1988 *Machiavellian Intelligence*. Oxford, UK: Oxford University Press.

10 Dunbar, R. I. M. 1998 The social brain hypothesis. *Evol Anthropol* **6**, 178-190. (DOI 10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8)

11 Whiten, A. & Byrne, R. W. 1988 Tactical deception in primates. *Behav Brain Sci* **11**, 233-273. (DOI 10.1017/S0140525X00049682)

12 Harcourt, A. . & Stewart, K. J. 1989 Functions of alliances in contests within wild gorilla groups. *Behaviour* **109**, 176-190. (DOI 10.1163/156853989X00213)

13 Dunbar, R. I. M. & Shultz, S. 2010 Bondedness and sociality. *Behaviour* **147**, 775-803. (DOI 10.1163/000579510X501151)

14 Silk, J. B. 2002 Using the’F’-word in primatology. *Behaviour* **139**, 421-446. (DOI 10.1163/156853902760102735)

15 Smuts, B. B. 1985 *Sex and friendship in baboons*. New York: Aldine Publishing Company.

16 Cheney, D. L. & Seyfarth, R. M. 2007 *Baboon metaphysics: the evolution of a social mind*. Chicago: University of Chicago Press.

17 Dunbar, R. I. M. 2009 The social brain hypothesis and its implications for social evolution. *Ann Hum Biol* **36**, 562-572. (DOI 10.1080/03014460902960289)

18 Silk, J. B., Alberts, S. C. & Altmann, J. 2003 Social bonds of female baboons enhance infant survival. *Science* **302**, 1231-1234. (DOI 10.1126/science.1088580)

19 Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., Wittig, R. M., Seyfarth, R. M. & Cheney, D. L. 2010 Strong and consistent social bonds enhance the longevity of female baboons. *Curr Biol* **20**, 1359-1361. (DOI 10.1016/j.cub.2010.05.067)

20 McComb, K. & Semple, S. 2005 Coevolution of vocal communication and sociality in primates. *Biol Lett* **1**, 381-385. (DOI 10.1098/rsbl.2005.0366)

21 Hauser, M. D. 1996 *The evolution of communication*. Cambridge, MA: MIT Press.

22 Mulligan, B. E. & Nellis, D. W. 1975 Vocal repertoire of the mongoose Herpestes Auropunctatus. *Behaviour* **55**, 237-257. (DOI 10.1163/156853975X00489)

23 Watts, C. H. S. 1975 Vocalizations of Australian hopping mice (Rodentia: Nutumys). *J Zool Lond* **177**, 247-263. (DOI 10.1111/j.1469-7998.1975.tb05982.x)

24 le Roux, A., Cherry, M. I. & Manser, M. B. 2009 The vocal repertoire in a solitary foraging carnivore, Cynictis penicillata, may reflect facultative sociality. *Naturwissenschaften* **96**, 575-584. (DOI 10.1007/s00114-008-0506-5)

25 Wong, J., Stewart, P. & Macdonald, D. 1999 Vocal repertoire in the European badger (Meles meles): Structure, context, and function. *J Mammal* **80**, 570-588. (DOI 10.2307/1383302)

26 Pollard, K. A. & Blumstein, D. T. 2011 Social group size predicts the evolution of individuality. *Curr Biol* **21**, 413-417. (DOI 10.1016/j.cub.2011.01.051)

27 Brady, C. A. 1981 The vocal repertoires of the bush dog (Speothos venaticus), crab-eating fox (Cerdocyon thous), and maned wolf (Chrysocyon brachyurus). *Anim Behav* **29**, 649-669. (DOI 10.1016/S0003-3472(81)80001-2)

28 Green, S. 1975 Communication by a graded vocal system in Japanese monkeys. In *Primate Behavior*, pp. 1-102. New York: Academic Press.

29 Leger, D. W., Owings, D. H. & Gelfand, D. L. 1980 Single-note vocalizations of California ground squirrels: graded signals and situation-specificity of predator and socially evoked calls. *Z Tierpsychol* **52**, 227-246. (DOI 10.1111/j.1439-0310.1980.tb00714.x)

30 Meise, K., Keller, C., Cowlishaw, G. & Fischer, J. 2011 Sources of acoustic variation: implications for production specificity and call categorization in chacma baboon (Papio ursinus) grunts. *J Acoust Soc Am* **129**, 1631-1641. (DOI 10.1121/1.3531944)

31 Peters, G. & A. Sliwa 1997 Acoustic communication in the aardwolf, Proteles cristatus (Carnivora: Hyaenidae). *Z Säugetierkd* **62**, 219-238.

32 Fischer, J., Metz, M., Seyfarth, R. M. & Cheney, D. L. 2001 Baboon responses to graded bark variants. *Anim Behav* **61**, 925-931. (DOI 10.1006/anbe.2000.1687)

33 Harris, M. A., Murie, J. O. & Duncan, J. A. 1983 Responses of Columbian ground squirrels to playback of recorded calls. *Z Tierpsychol* **63**, 318-330. (DOI 10.1111/j.1439-0310.1983.tb00747.x)

34 le Roux, A., Jackson, T. P. & Cherry, M. I. 2001 Does Brants' whistling rat (Parotomys brantsii) use an ugency-based alarm system in reaction to aerial and terrestrial predators? *Behaviour* **138**, 757-773. (DOI 10.1163/156853901752233398)

35 Ackers, S. H. & Slobodchikoff, C. N. 1999 Communication of stimulus size and shape in alarm calls of gunnison's prairie dogs, Cynomys gunnisoni. *Ethology* **105**, 149-162. (DOI 10.1046/j.1439-0310.1999.00381.x)

36 Greene, G. & Meagher, T. 1998 Red squirrels, Tamiasciurus hudsonicus, produce predator-class specific alarm calls. *Anim Behav* **55**, 511-518. (DOI 10.1006/anbe.1997.0620)

37 Manser, M. B. 2001 The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proc Biol Sci* **268**, 2315-2324. (DOI 10.1098/rspb.2001.1773)

38 Crockford, C. & Boesch, C. 2003 Context-specific calls in wild chimpanzees, Pan troglodytes verus: analysis of barks. *Anim Behav* **66**, 115-125. (DOI 10.1006/anbe.2003.2166)

39 Seyfarth, R. M., Cheney, D. L. & Marler, P. 1980 Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Anim Behav* **28**, 1070-1094. (DOI 10.1016/S0003-3472(80)80097-2)

40 Stephan, C. & Zuberbühler, K. 2008 Predation increases acoustic complexity in primate alarm calls. *Biol Lett* **4**, 641-644. (DOI 10.1098/rsbl.2008.0488)

41 Macedonia, J. M. & Evans, C. S. 1993 Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* **93**, 177-197. (DOI 10.1111/j.1439-0310.1993.tb00988.x)

42 Hauser, M. D., Teixidor, P., Field, L. & Flaherty, R. 1993 Food-elicited calls in chimpanzees - effects of food quantity and divisability. *Anim Behav* **45**, 817-819. (DOI 10.1006/anbe.1993.1096)

43 Roush, R. S. & Snowdon, C. T. 2000 Quality, quantity, distribution and audience effects on food calling in cotton—top tamarins. *Ethology* **106**, 673-690. (DOI 10.1046/j.1439-0310.2000.00581.x)

44 Ouattara, K., Lemasson, A. & Zuberbühler, K. 2009 Campbell’s monkeys concatenate vocalizations into context-specific call sequences. *Proc Natl Acad Sci U S A* **106**, 22026-22031. (DOI 10.1073/pnas.0908118106)

45 Arnold, K. & Zuberbühler, K. 2006 The alarm-calling system of adult male putty-nosed monkeys, Cercopithecus nictitans martini. *Anim Behav* **72**, 643-653. (DOI 10.1016/j.anbehav.2005.11.017)

46 Cap, H., Deleporte, P., Joachim, J. & Reby, D. 2008 Male vocal behavior and phylogeny in deer. *Cladistics* **24**, 917-931. (DOI 10.1111/j.1096-0031.2008.00223.x)

47 Knornschild, M., Glockner, V. & von Helversen, O. 2010 The vocal repertoire of two sympatric species of nectar-feeding bats (Glossophaga soricina and G. commissarisi). *Acta Chiropterol* **12**, 205-215. (DOI 10.3161/150811010X504707)

48 Palombit, R. A. 1992 A preliminary study of vocal communication in wild long-tailed macaques (Macaca fascicularis). 1. Vocal repertoire and call emission. *Int J Primatol* **13**, 143-182. (DOI 10.1007/BF02547839)

49 Masataka, N. 1982 A field study on the vocalizations of Goeldi’s monkeys (Callimico goeldii). *Primates* **23**, 206-219. (DOI 10.1007/BF02381161)

50 Goodall, J. 1986 *The chimpanzees of Gombe*. Cambridge, MA: Belknapp.

51 Sieber, O. J. 1984 Vocal communication in raccoons (Procyon lotor). *Behaviour* **90**, 80-113. (DOI 10.1163/156853984X00560)

52 Robbins, R. L. 2000 Vocal communication in free-ranging African wild dogs (Lycaon pictus). *Behaviour* **137**, 1271-1298. (DOI 10.1163/156853900501926)

53 Hamilton, 3rd, W.J. & Arrowood, P. C. 1978 Copulatory vocalizations of chacma baboons (Papio ursinus), gibbons (Hylobates hoolock), and humans. *Science* **200**, 1405-1409. (DOI 10.1126/science.663622)

54 Henzi, P. S. 1996 Copulation calls and paternity in chacma baboons. *Anim Behav* **51**, 233-234. (DOI 10.1006/anbe.1996.0021)

55 Gros-Louis, J. 2002 Contexts and behavioral correlates of trill vocalizations in wild white-faced capuchin monkeys (Cebus capucinus). *Am J Primatol* **57**, 189-202. (DOI 10.1002/ajp.10042)

56 Silk, J. B., Cheney, D. L. & Seyfarth, R. M. 1996 The form and function of post-conflict interactions between female baboons. *Anim Behav* **52**, 259-268. (DOI 10.1006/anbe.1996.0171)

57 Palombit, R. A., Cheney, D. L. & Seyfarth, R. M. 1999 Male grunts as mediators of social interaction with females in wild chacma baboons (Papio cynocephalus ursinus). *Behaviour* **136**, 221-242. (DOI 10.1163/156853999501298)

58 Peters, G. & Tonkin-Leyhausen, B. A. 1999 Evolution of acoustic communication signals of mammals: friendly close-range vocalizations in Felidae (Carnivora). *J Mamm Evol* **6**, 129-159. (DOI 10.1023/A:1020620121416)

59 Koda, H., Shimooka, Y. & Sugiura, H. 2008 Effects of caller activity and habitat visibility on contact call rate of wild Japanese macaques (Macaca fuscata). *Am J Primatol* **70**, 1055-1063. (DOI 10.1002/ajp.20597)

60 Manser, M. B. 1998 The evolution of auditory communication in suricates, Suricata suricatta.

61 Marler, P. 1977 The evolution of communication. In *How animals communicate* (ed T. A. Sebeok), pp. 45-70. Bloomington, IN: Indiana University Press.

62 Marler, P. & Mitani, J. 1988 Vocal communication in primates and birds: parallels and contrasts. In *Primate vocal communication* (eds D. Todt P. Goedeking & D. Symmes), pp. 3-14. Berlin: Springer.

63 Philips, M. & Austad, S. N. 1990 Animal communication and social evolution. In *Interpretation and explanation in the study of animal behavior. Vol. 1. Interpretation, intentionality and communication* (eds M. Bekoff & D. Jamieson), pp. 254-268. Boulder, CO: Westview.

64 Waser, P. M. 1982 The evolution of male loud calls among mangabeys and baboons. In *Primate communication* (eds C. T. Snowdon C. H. Brown & M. R. Petersen), pp. 117-143. Cambridge, UK: Cambridge University Press.

65 Janik, V. M. & Slater, P. J. B. 1997 Vocal learning in mammals. *Adv Stud Behav* **26**, 59-99. (DOI 10.1016/S0065-3454(08)60377-0)

66 Zuberbühler, K. 2003 Referential signaling in non-human primates: Cognitive precursors and limitations for the evolution of language. *Adv Stud Behav* **33**, 265-307. (DOI 10.1016/S0065-3454(03)33006-2)

67 Seyfarth, R. M. & Cheney, D. L. 2010 Production, usage, and comprehension in animal vocalizations. *Brain Lang* **115**, 92-100. (DOI 10.1016/j.bandl.2009.10.003)

68 Fant, G. 1960 *Acoustic theory of speech production. With calculations based on X-ray studies of Russian articulations.* 's-Gravenhage: Mouton.

69 Fitch, W. T. 2000 The evolution of speech: a comparative review. *Trends Cogn Sci* **4**, 258-267. (DOI 10.1016/S1364-6613(00)01494-7)

70 Hauser, M. D. 1992 Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: a learned mode of production? *J Acoust Soc Am* **91**, 2175-2179. (DOI 10.1121/1.403676)

71 Rendall, D., Rodman, P. S. & Emond, R. E. 1996 Vocal recognition of individuals and kin in free-ranging rhesus monkeys. *Anim Behav* **51**, 1007-1015. (DOI 10.1006/anbe.1996.0103)

72 Fitch, W. T. 1997 Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *J Acoust Soc Am* **102**, 1213-1222. (DOI 10.1121/1.421048)

73 Fitch, W. T. & Giedd, J. 1999 Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J Acoust Soc Am* **106**, 1511-1522. (DOI 10.1121/1.427148)

74 Fitch, W. T. & Kelley, J. P. 2000 Perception of Vocal Tract Resonances by Whooping Cranes Grus americana. *Ethology* **106**, 559-574. (DOI 10.1046/j.1439-0310.2000.00572.x)

75 Lieberman, P. H., Klatt, D. H. & Wilson, W. H. 1969 Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science* **164**, 1185-1187. (DOI 10.1126/science.164.3884.1185)

76 Snowdon, C. T. 2009 Plasticity of communication in nonhuman primates. *Adv Stud Behav* **40**, 239-276. (DOI 10.1016/S0065-3454(09)40007-X)

77 Pepperberg, I. M. 1981 Functional vocalizations by an African Grey parrot (Psittacus erithacus). *Z Tierpsychol* **55**, 139-160. (DOI 10.1111/j.1439-0310.1981.tb01265.x)

78 Ralls, K., Fiorelli, P. & Gish, S. 1985 Vocalizations and vocal mimicry in captive harbor seals, Phoca vitulina. *Can J Zool* **63**, 1050-1056. (DOI 10.1139/z85-157)

79 Konrad, R. & Geissmann, T. 2006 Vocal diversity and taxonomy of Nomascus in Cambodia. *Int J Primatol* **27**, 713-745. (DOI 10.1007/s10764-006-9042-3)

80 Méndez-Cárdenas, M., Randrianambinina, B., Rabesandratana, A., Rasoloharijaona, S. & Zimmermann, E. 2008 Geographic variation in loud calls of sportive lemurs (Lepilemur ssp.) and their implications for conservation. *Am J Primatol* **70**, 828-838. (DOI 10.1002/ajp.20554)

81 Mitani, J. C., Hunley, K. L. & Murdoch, M. E. 1999 Geographic variation in the calls of wild chimpanzees: A reassessment. *Am J Primatol* **47**, 133-151. (DOI 10.1002/(SICI)1098-2345(1999)47:2<133::AID-AJP4>3.0.CO;2-I)

82 Moore, J. L., Manne, L., Brooks, T., Burgess, N. D., Davies, R., Rahbek, C., Williams, P. & Balmford, A. 2002 The distribution of cultural and biological diversity in Africa. *Proc Biol Sci* **269**, 1645-1653. (DOI 10.1098/rspb.2002.2075)

83 Slocombe, K. E. & Zuberbühler, K. 2007 Chimpanzees modify recruitment screams as a function of audience composition. *Proc Natl Acad Sci U S A* **104**, 17228-17233. (DOI 10.1073/pnas.0706741104)

84 Slocombe, K. E., Kaller, T., Turman, L., Townsend, S. W., Papworth, S., Squibbs, P. & Zuberbühler, K. 2010 Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. *Behav Ecol Sociobiol* **64**, 1959-1966. (DOI 10.1007/s00265-010-1006-0)

85 Papworth, S., Böse, A., Barker, J., Schel, A. M. & Zuberbühler, K. 2008 Male blue monkeys alarm call in response to danger experienced by others. *Biol Lett* **4**, 472-475. (DOI 10.1098/rsbl.2008.0299)

86 Di Bitetti, M. S. 2005 Food-associated calls and audience effects in tufted capuchin monkeys, Cebus apella nigritus. *Anim Behav* **69**, 911-919. (DOI 10.1016/j.anbehav.2004.05.021)

87 Pollick, A. S., Gouzoules, H. & de Waal, F. B. M. 2005 Audience effects on food calls in captive brown capuchin monkeys, Cebus apella. *Anim Behav* **70**, 1273-1281. (DOI 10.1016/j.anbehav.2005.03.007)

88 Zuberbühler, K., Cheney, D. L. & Seyfarth, R. M. 1999 Conceptual semantics in a nonhuman primate. *J Comp Psychol* **113**, 33-42. (DOI 10.1037/0735-7036.113.1.33)

89 Doupe, A. J. & Kuhl, P. K. 1999 Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci* **22**, 567-631. (DOI 10.1146/annurev.neuro.22.1.567)

90 Bouchet, H., Pellier, A., Blois-Heulin, C. & Lemasson, A. 2010 Sex differences in the vocal repertoire of adult red-capped mangabeys (Cercocebus torquatus): A multi-level acoustic analysis. *Am J Primatol* **72**, 360-375. (DOI 10.1002/ajp.20791)

91 Bohn, K. M., Schmidt-French, B., Schwartz, C., Smotherman, M. & Pollak, G. D. 2009 Versatility and stereotypy of free-tailed bat songs. *PloS One* **4**, 6746. (DOI 10.1371/journal.pone.0006746)

92 Hockett, C. F. 1960 The origin of speech. *Sci Am* **203**, 88-96. (DOI 10.1038/scientificamerican0960-88)

93 Charles-Dominique, P. 1977 *Ecology and behaviour of nocturnal primates: prosimians of equatorial West Africa*. New York: Columbia University Press.

94 Cleveland, J. & Snowdon, C. T. 1982 The complex vocal repertoire of the adult cotton-top tamarin (Saguinus oedipus oedipus). *Z Tierpsychol* **58**, 231-270. (DOI 10.1111/j.1439-0310.1982.tb00320.x)

95 Hohmann, G. 1991 Comparative analyses of age-specific and sex-specific patterns of vocal behavior in 4 species of Old-world monkeys. *Folia Primatol* **56**, 133-156. (DOI 10.1159/000156538)

96 MacLanahan, E. B. & Green, K. M. 1977 The vocal repertoire and an analysis of the contexts of vocalizations in Leontopithecus rosalia. In *The biology and conservation of the Callitrichidae* (ed D. G. Kleiman), pp. 251-269. Washington, DC: Smithsonian Institute.

97 Kudo, H. 1987 The study of vocal communication of wild mandrills in Cameroon in relation to their social structure. *Primates* **28**, 289-308. (DOI 10.1007/BF02381013)

98 Aiello, L. C. & Dunbar, R. I. M. 1993 Neocortex size, group size, and the evolution of language. *Curr Anthropol* **34**, 184-193. (DOI 10.1086/204160)

99 Dunbar, R. I. M. & Dunbar, P. 1975 *Social dynamics of gelada baboons*. Basel, New York: Karger.

100 Richman, B. 1987 Rhythm and melody in gelada vocal exchanges. *Primates* **28**, 199-223. (DOI 10.1007/BF02382570)

101 Page, S. L., Chiu, C. & Goodman, M. 1999 Molecular Phylogeny of Old World Monkeys (Cercopithecidae) as Inferred from γ-Globin DNA Sequences. *Mol Phylogenet Evol* **13**, 348-359. (DOI 10.1006/mpev.1999.0653)

102 Bergman, T. J. & Kitchen, D. M. 2009 Comparing responses to novel objects in wild baboons (Papio ursinus) and geladas (Theropithecus gelada). *Anim Cogn* **12**, 63-73. (DOI 10.1007/s10071-008-0171-2)

103 Beehner, J. C., Bergman, T. J., Cheney, D. L., Seyfarth, R. M. & Whitten, P. L. 2006 Testosterone predicts future dominance rank and mating activity among male chacma baboons. *Behav Ecol Sociobiol* **59**, 469-479. (DOI 10.1007/s00265-005-0071-2)

104 le Roux, A., Beehner, J. C. & Bergman, T. J. 2011 Female philopatry and dominance patterns in wild geladas. *Am J Primatol* **73**, 422-430. (DOI 10.1002/ajp.20916)

105 Snyder-Mackler, N., Bergman, T. J. & Beehner, J. C. In Press Defining higher levels in a gelada multilevel society. *Int J Primatol*

106 Kawai, M., Ohsawa, H., Mori, U. & Dunbar, R. 1983 Social organization of gelada baboons: social units and definitions. *Primates* **24**, 13-24. (DOI I: 10.1007/BF02381450)

107 Bergman, T. J. 2010 Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proc Biol Sci* **277**, 3045-3053. (DOI 10.1098/rspb.2010.0580)

108 le Roux, A. & Bergman, T. J. In Press Indirect rival assessment in a social primate, Theropithecus gelada. *Anim Behav*

109 Dunbar, R. I. M. 1984 *Reproductive decisions: an economic analysis of gelada baboon social strategies*. Princeton, NJ: Princeton University Press.

110 Mori, U. 1979 Reproductive behaviour. In *Ecological and sociological studies of gelada baboons*, pp. 183-199. Tokyo: Kodansha.

111 Dunbar, R. I. M. 1977 Feeding ecology of gelada baboons: a preliminary report. In *Primate ecology* (ed T. H. Clutton-Brock), pp. 250-273. London: Academic Press.

112 Dunbar, R. I. M. & Bose, U. 1991 Adaptation to grass-eating in gelada baboons. *Primates* **32**, 1-7. (DOI 10.1007/BF02381596)

113 Iwamoto, T. 1993 The ecology of Theropithecus gelada. In *Theropithecus: the rise and fall of a primate genus* (ed N. G. Jablonski), pp. 441-452. Cambridge, UK: Cambridge University Press.

114 Fischer, J., Hammerschmidt, K., Cheney, D. L. & Seyfarth, R. M. 2001 Acoustic features of female chacma baboon barks. *Ethology* **107**, 33-54.

115 Owren, M. J., Seyfarth, R. M. & Cheney, D. L. 1997 The acoustic features of vowel-like grunt calls in chacma baboons (Papio cyncephalus ursinus): implications for production processes and functions. *J Acoust Soc Am* **101**, 2951-2963. (DOI 10.1121/1.418523)

116 Rendall, D., Seyfarth, R. M., Cheney, D. L. & Owren, M. J. 1999 The meaning and function of grunt variants in baboons. *Anim Behav* **57**, 583-592. (DOI 10.1006/anbe.1998.1031)

117 Fischer, J., Hammerschmidt, K., Cheney, D. L. & Seyfarth, R. M. 2002 Acoustic features of male baboon loud calls: Influences of context, age, and individuality. *J Acoust Soc Am* **111**, 1465-1974. (DOI 10.1121/1.1433807)

118 Altmann, S. A. 1998 *Foraging for survival: Yearling baboons in Africa*. Chicago: University of Chicago Press.

119 Altmann, S. A. & Altmann, J. 1970 *Baboon ecology*. Chicago: University of Chicago Press.

120 Norton, G. W., Rhine, R. J., Wynn, G. W. & Wynn, R. D. 1987 Baboon diet: a five-year study of stability and variability in the plant feeding and habitat of the yellow baboon (Papio cynocephalus) of Mikumi National Park, Tanzania. *Folia Primatol* **48**, 78-120. (DOI 10.1159/000156287)

121 Rowell, T. E. 1966 Forest living baboons in Uganda. *J Zool* **149**, 344-364. (DOI 10.1111/j.1469-7998.1966.tb04054.x)

122 Hall, K. R. L. & DeVore, I. 1965 Baboon social behavior. In *Primate Social Behavior* (eds I. DeVore & K. R. L. Hall), pp. 53-110. New York: Holt, Rinehart & Winston.

123 Aich, H., Moos-Heilen, R. & Zimmermann, E. 1990 Vocalizations of adult gelada baboons (Theropithecus gelada): acoustic structure and behavioural context. *Folia Primatol* **55**, 109-132. (DOI 10.1159/000156508)

124 Lehner, P. N. 1996 *Handbook of ethological methods*. Cambridge, NY: Cambridge University Press.

125 Dunbar, R. I. M. 1993 Coevolution of neocortical size, group size and language in humans. *Behav Brain Sci* **16**, 681-693. (DOI 10.1017/S0140525X00032325)

126 Schreier, A. L. & Swedell, L. 2009 The fourth level of social structure in a multi-level society: ecological and social functions of clans in hamadryas baboons. *Am J Primatol* **71**, 948-955. (DOI 10.1002/ajp.20736)

127 Aureli, F., Preston, S. D. & de Waal, F. 1999 Heart rate responses to social interactions in free-moving rhesus macaques (Macaca mulatta): A pilot study. *J Comp Psychol* **113**, 59-65. (DOI 10.1037/0735-7036.113.1.59)

128 Dunbar, R. I. M. 2010 The social role of touch in humans and primates: behavioural function and neurobiological mechanisms. *Neurosci Biobehav Rev* **34**, 260-268. (DOI 10.1016/j.neubiorev.2008.07.001)

129 Keverne, E. B., Martensz, N. D. & Tuite, B. 1989 Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology* **14**, 155-161. (DOI 10.1016/0306-4530(89)90065-6)

130 Baldwin, J. D. & Baldwin, J. I. 1976 Vocalizations of howler monkeys (Alouatta palliata) in southwestern Panama. *Folia Primatol* **26**, 81-108. (DOI 10.1159/000155733)

131 Robinson, J. G. 1979 An analysis of the organization of vocal communication in the titi monkey Callicebus moloch. *Z Tierpsychol* **49**, 381-405. (DOI 10.1111/j.1439-0310.1979.tb00300.x)

132 Robinson, J. G. 1984 Syntactic structures in the vocalizations of wedge-capped capuchin monkeys, Cebus olivaceus. *Behaviour* **90**, 46-79. (DOI 10.1163/156853984X00551)

133 Range, F. & Fischer, J. 2004 Vocal repertoire of sooty mangabeys (Cercocebus torquatus atys) in the Taï National Park. *Ethology* **110**, 301-321. (DOI 10.1111/j.1439-0310.2004.00973.x)

134 Strushaker, T. T. 1967 Auditory communication among vervet monkeys (Cercopithecus aethiops). In *Social communication among primates* (ed S. A. Altmann), Chicago: University of Chicago Press.

135 Bermejo, M. & Omedes, A. 1999 Preliminary vocal repertoire and vocal communication of wild bonobos (Pan paniscus) at Lilungu (Democratic Republic of Congo). *Folia Primatol* **70**, 328-357. (DOI 10.1159/000021717)

136 MacKinnon, J. 1974 The behaviour and ecology of wild orangutans (Pongo pygmaeus). *Anim Behav* **22**, 3-74. (DOI 10.1016/S0003-3472(74)80054-0)

137 Strushaker, T. T. 1975 *The red colobus monkey*. Chicago: University of Chicago Press.

138 Long, C. V. 2007 Vocalisations of the degu Octodon degus, a social caviomorph rodent. *Bioacoustics* **16**, 223-244.

139 Cheney, D. L., Seyfarth, R. M. & Silk, J. B. 1995 The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Anim Behav* **50**, 249-257. (DOI 10.1006/anbe.1995.0237)

140 Richman, B. 1976 Some vocal distinctive features used by gelada monkeys. *J Acoust Soc Am* **60**, 718-776. (DOI 10.1121/1.381144)

141 Fischer, J., Kitchen, D. M., Seyfarth, R. M. & Cheney, D. L. 2004 Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. *Behav Ecol Sociobiol* **56**, 140-148. (DOI 10.1007/s00265-003-0739-4)

142 Kitchen, D. M., Cheney, D. L. & Seyfarth, R. M. 2005 Male chacma baboons (Papio hamadryas ursinus) discriminate loud call contests between rivals of different relative ranks. *Anim Cogn* **8**, 1-6. (DOI 10.1007/s10071-004-0222-2)

143 Wittig, R. M., Crockford, C., Seyfarth, R. M. & Cheney, D. L. 2007 Vocal alliances in Chacma baboons (Papio hamadryas ursinus). *Behav Ecol Sociobiol* **61**, 899-909. (DOI 10.1007/s00265-006-0319-5)

144 Andrew, R. J. 1976 Use of formants in the grunts of baboons and other nonhuman primates. *Ann N Y Acad Sci* **280**, 673-693. (DOI 10.1111/j.1749-6632.1976.tb25530.x)

145 Palagi, E., Leone, A., Mancini, G. & Ferrari, P. F. 2009 Contagious yawning in gelada baboons as a possible expression of empathy. *Proc Natl Acad Sci U S A* **106**, 19262-19267. (DOI 10.1073/pnas.0910891106)

Table 2.1. Vocal repertoire size for exemplar species from Primata, Rodentia, and Carnivora broken down into six categories: Allospecific (alarm calls and food calls), Long distance (separation calls, intergroup spacing calls), Contact (short-range soft calls), Competitive (threat and display calls), Distress calls (fear calls during agonism), and Other (contexts unknown or made in several different contexts).

Species name	Allospecific	Long distance	Contact	Competitive	Distress	Other	Total size	Citation
Order Primata								
Alouatta palliata	-	1	-	8	2	1	12	[130]
Arctocebus calabarensis	-	1	-	1	1	-	3	[93]
Callicebus moloch	1	3	1	1	3	1	10	[131]
Callimico goeldii	4	7	3	6	4	1	25	[49]
Cebus olivaceus	-	-	4	4	2	1	11	[132]
Cercocebus torquatus	4	1	3	5	1	-	14	[133]
Cercopithecus aethiops	5	-	3	5	3	3	19	[134]
Euoticus elegantulus	2	1	2	-	1	-	6	[93]
Galago alleni	1	-	2	1	1	-	5	[93]
G. demidovii	1	1	3	1	1	1	8	[93]
Macaca fascicularis	2	2	1	1	4	5	15	[48]
M. radiata	1	1	3	7	4	3	19	[95]
M. silenus	1	2	3	5	2	1	14	[95]
Mandrillus sphinx	1	3	1	-	4	-	9	[97]
Pan paniscus	-	-	4	-	1	4	9	[135]
P. troglodytes	3	1	8	4	8	3	27	[50]
Perodicticus potto	-	1		2	1	1	5	[93]
Pongo pygmaeus	1	1	1	2	2	1	8	[136]
Presbytis entellus	3	1	2	3	3	2	14	[95]
P. johnii	2	1	3	4	3	1	14	[95]
Procolobus badius	2	-	-	5	2	2	11	[137]
Order Rodentia								
Notomys alexis	-	-	1	1	1	2	5	[23]
N. cervinus	-	-	-	1	1	2	4	[23]
N. mitchellii	-	-	1	1	1	2	5	[23]
N. fuscus	-	-	1	1	1	2	5	[23]
Octodon degus	2	-	4	5	2	-	13	[138]

Order Carnivora								
Lycaon pictus	3	2	7	5	2	6	25	[52]
Suricata suricatta	11	-	7	2	2	3	25	[37,60]
Cynictis penicillata	3	-	2	2	-	1	8	[24]
Speothos venaticus	-	1	3	2	1	-	7	[27]
Cerdocyon thous	2	1	2	1	-	-	6	[27]
Chrysocyon brachyurus	1	2	2	2	1	-	8	[27]

Note: Sources from Primata are drawn from the repertoire analysis made by McComb and Semple [20], excluding captive studies. Total repertoire sizes in this paper are slightly different because we did not count sequences of discrete call units as separate calls if the units were produced singly. Sounds that are not strictly ‘vocalisations’, such as sneezes, coughs, and teeth chattering, are excluded from the table. For comparison, we focus on exemplar species from Rodentia and Carnivora because of similarities in social and vocal behaviour.

Table 2.2. Descriptions of call types used by geladas and chacma baboons in short-range and in long-distance situations, including the way in which they are physically produced and the contexts in which they occur. CF=chacma female; CM=chacma male; GF=gelada female; GM=gelada male. Asterisks denote vocalisations that were not used in discriminant function analyses due to low sample size.

I. Shared vocalisations in chacma baboons and geladas

Call type	Mode of production	Context
affiliative grunt (CF, CM, GF, GM)	exhale	A soft tonal contact call used during approaching, grooming, and infant-handling, as well as while moving and foraging [30,100,123,139,140]
copulation call (CF, CM*, GF, GM)	exhale	Loud grunts given before and during mating [123]
fear bark (CF, CM*, GF)	exhale with retracted lips	A 'cough-like' vocalisation [141] given by subordinate individuals to high-ranking animals [123]
threat grunt (CF, CM, GF, GM)	exhale	A staccato-like vocalisation uttered by the dominant individual in an aggressive encounter [123,142,143]
alarm call (CF*, CM*, GF)	exhale	Noisy, harsh calls used in response to predators and other environmental threats [100,114]
display call or 'wahoo' (CM, GF*, GM)	inhale and exhale	Loud calls typically uttered during competitive displays. [123,141]. Chacma and gelada males, in particular, make a ‘roar’ that often comes before these wahoo calls.
lost call (CM*,CF*,GF*,GM*)	long exhale	A noisy vocalisation that rises in pitch towards the end of the call and associated with separation from the group or particular individuals
scream (CF, GF, GM)	long exhale with retracted lips	A noisy drawn-out defensive call, usually given by subordinates when attacked by a higher ranking individual [100,123,139]

II. Derived gelada vocalisations

Call type	Mode of production	Context
inhaled grunt (GF*, GM)	vocal inhale	Vocalised inhales often part of an affiliative grunt calling bout [140]. Sometimes, inhaled grunts can have an audibly ‘nasal’ sound, produced by the withdrawn lip obscuring nasal passages [140].
moan (GF*, GM)	long exhale (sometimes inhaled)	Long drawn-out affiliative grunt, often given by leader males to their unit’s females ([100,123,144]; this study)

wobble (GF*, GM)	vocal inhale or exhale with lip or tongue-flicking	Soft, undulating calls usually given by males to their unit females, often when nervous (this study).
yawn (GF*, GM)	inhale	A vocalised yawn given in social contexts, often involving grooming sessions and also after mating or in competitive situations ([145]; this study)
pre-copulation call (GF)	short exhale	Calls given by estrous females while presenting their genitals to males
how barks (GM*)	exhale	High-pitched barks/ whinnies given by non-leader males giving chase to other males in competitive displays

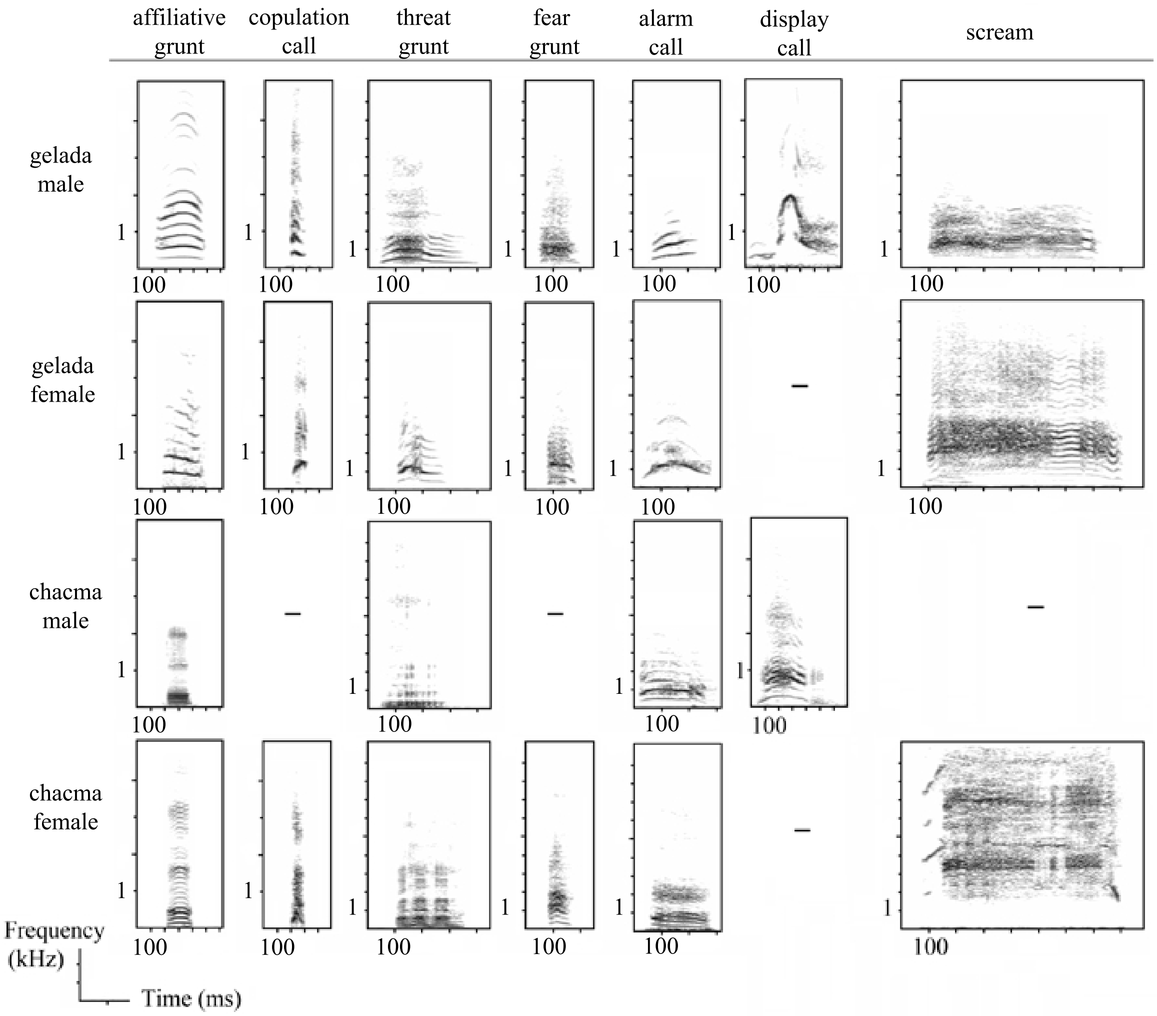


Figure 2.1. Spectrograms of homologous calls shared by geladas and chacma baboons. Dashes represent calls that were not produced or produced at a very low rate.

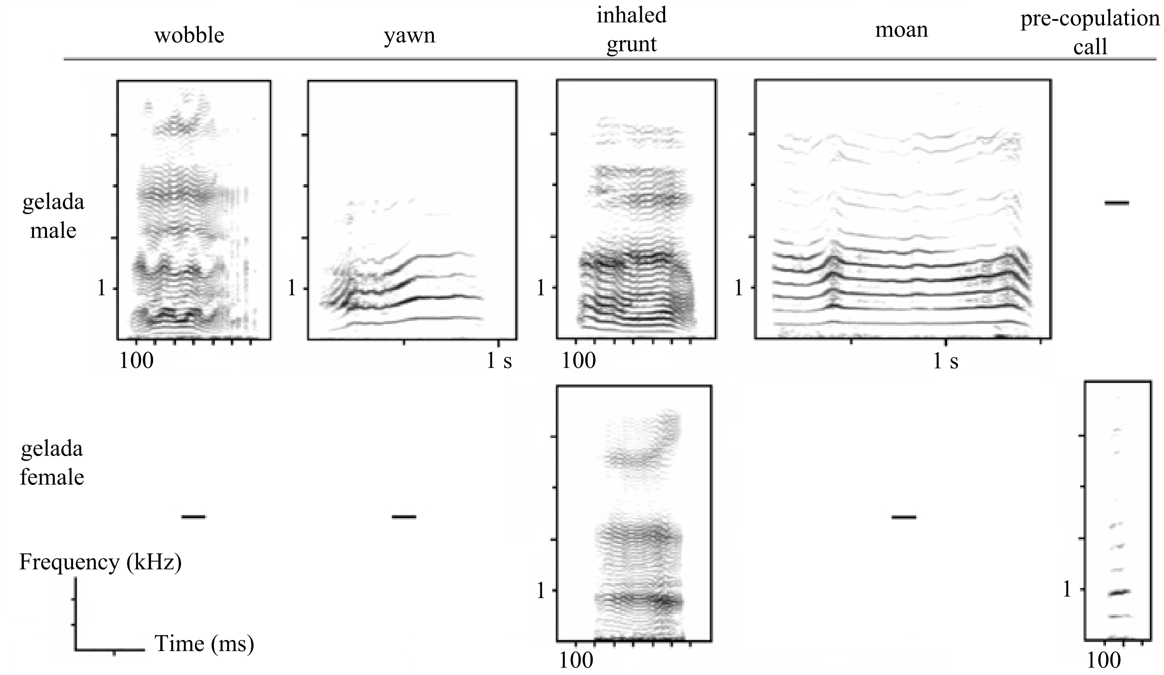


Figure 2.2. Spectrograms of derived call types produced only by geladas. Dashes represent calls that were not produced or produced at a very low rate.

Chapter 3:

Divergent acoustic properties of gelada and baboon vocalizations and implications for the evolution of human speech

Abstract

Human speech has many complex spectral and temporal features traditionally thought to be absent in the vocalizations of other primates. Recent explorations of the vocal capabilities of non-human primates are challenging this view. Here, we continue this trend by exploring the spectro-temporal properties of gelada (*Theropithecus gelada*) vocalizations. First, we made cross-species comparisons of geladas, chacma baboons, and human vowel space area. We found that adult male and female gelada exhaled grunts – a call type shared with baboons – have formant profiles that overlap more with human vowel space than do baboon grunts. These gelada grunts also contained more modulation of fundamental and formant frequencies than did baboon grunts. Second, we compared formant profiles and modulation of exhaled grunts to the derived call types (those not shared with baboons) produced by gelada males. These derived calls contained divergent formant profiles, and a subset of them, notably wobbles and vocalized yawns, were more modulated than grunts. Third, we investigated the rhythmic patterns of wobbles, a call type shown previously to contain cycles that match the 3-8 Hz tempo of speech. We use a larger dataset to show that the wobble rhythm overlaps more with speech rhythm than previously thought. We also found that variation in cycle duration depends on the production modality; specifically, exhaled wobbles were produced at a slower tempo than inhaled wobbles. Moreover, the variability in cycle duration within wobbles aligns with a linguistic property known as "Menzerath's law" in that there was a negative association between cycle duration and wobble size (i.e., the number of cycles). Taken together, our results add to growing evidence that non-human primates are anatomically capable of producing modulated sounds. Specifically, gelada vocalizations – as compared to baboons – appear to have evolved spectro-temporal characteristics that are more akin to the acoustic properties of human speech. These findings support and expand on current hypotheses of speech evolution, including the "neural hypothesis" and the "bimodal speech rhythm hypothesis".

1. Introduction

Human speech is a complex trait encompassing both spectral and temporal features that are argued to be unique among primates (Fitch, 2000, 2010; Ghazanfar, 2013). Knowing the extent to which these 'unique' acoustic features of speech are due to special physical and mechanical adaptations (in addition to a behavioral and neurobiological ones), however, requires a clear understanding of the physical limitations of our non-human primate relatives. To better differentiate the features that are truly unique to human speech from those which are shared, explorations of the vocal limitations of monkeys and apes are becoming increasingly common (Boë et al., 2017; Fitch, de Boer, Mathur, & Ghazanfar, 2016). Here, we summarize contemporary views on spectral and temporal features thought to be evolutionarily derived – or 'unique' – acoustic features of speech compared to the vocal capacities of other primates. For each of these unique acoustic features, we briefly review current findings from non-human primate studies that put the ‘uniqueness’ of each feature into question. Additionally, we further test the boundaries of what makes human speech unique by providing new analyses of the spectro-temporal vocal capacities of the gelada (*Theropithecus gelada*), a monkey known for its dynamic vocal behavior (Aich, Moos-Heilen, & Zimmermann, 1990; Benítez, le Roux, Fischer, Beehner, & Bergman, 2016; Bergman, 2013; Gustison, le Roux, & Bergman, 2012; Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016; Richman, 1976, 1987).

1.1 Spectral features of vocal production in humans and other primates

Source-filter theory is a well-established framework used to understand spectral components of vocal signal production in humans and other terrestrial vertebrates (Chiba & Kajiyama, 1941; Fant, 1960; Fitch & Suthers, 2016; Taylor, Charlton, & Reby, 2016; Taylor & Reby, 2010). This theory describes vocal signal production as a two-step process in which specific parts of the vocal apparatus contribute in independent ways to the final vocal signal. The “source” signal, or “glottal wave”, is created by vibrations in the vocal folds that result in a series of frequency components known as the fundamental frequency (f0) and its harmonic overtones (Titze et al., 2015). The “filter” signal, is the result of the source signal being molded by the resonance properties of the supralaryngeal vocal tract. These resonance properties are dictated by the shape of the oral and nasal cavities which results in emphasized frequencies, or “formants”, denoted by F_n (Titze et al., 2015). Humans appear to be the only primates able to flexibly control the formants of their vocal signals, primarily through the modification of articulators (e.g., tongue, lips, velum, and lower jaw); the relationships between the first three formants make up distinct vowels (Fitch, 2000; Ghazanfar & Rendall, 2008; Lieberman, Klatt, & Wilson, 1969). The five monophthong vowels most common across languages are /ɑ/, /i/, /ɛ /, /o/, and /u/, and many of these make up the extreme corners of vowel space area (VSA) (Maddieson, 1984). English VSA, for example, ranges between /i/ as in 'beet' with a low-frequency F1 and high-frequency F2, /u/ as in 'boot' with a low-frequency F1 and a low-frequency F2, and /a/ or /ɑ/ as in 'boss' with a high-frequency F1 and a moderate-frequency F2 (Fitch & Hauser, 1995).

A widely accepted hypothesis about why non-human primates lack control of their formant profiles is that they are limited by the anatomical design of their vocal tracts. Referred to as the "peripheral hypothesis" (Fitch et al., 2016), this hypothesis originally gained traction in the 1960s when Lieberman and colleagues (Lieberman et al., 1969) simulated the formants of a rhesus macaque based on vocal tract shape variation in an anesthetized monkey to conclude that rhesus macaques have a physically constrained range of acoustic variability. This hypothesis was supported by a similar study on chimpanzees (Lieberman, Crelin, & Klatt, 1972) and remains the prevailing hypothesis to explain why humans appear to have the unique ability to flexibly modulate the spectral properties of their vocalizations (D. Crystal, 2003; Raphael, Borden, & Harris, 2007; Yule, 2006).

An alternative hypothesis is that non-human primates have the anatomical capacity to modulate their vocal system in a speech-like way, but lack the brain mechanisms needed to do so (Hockett, 1960). Referred to as the "neural hypothesis" (Fitch et al., 2016), this hypothesis is gaining attention as emerging research on monkey and ape vocal systems challenge the peripheral hypothesis. Case studies on captive gorillas (*Gorilla gorilla*) and orangutans (*Pongo abelii*) suggest that great apes can control their vocal tract enough to acquire new vocalizations with vowel- or consonant-like properties (Lameira, Hardus, Mielke, Wich, & Shumaker, 2016; Lameira, Maddieson, & Zuberbühler, 2014; Perlman & Clark, 2015). Moreover, old world monkeys like chacma baboons (*Papio ursinus*), hamadryas baboons (*Papio hamadryas*), and rhesus macaques (*Macaca mulatta*) produce short vocalizations – often referred to as “grunts” – that have formant profiles that align closely with human vowels (Andrew, 1976; Boë et al., 2017; Fitch, 1997; Ghazanfar et al., 2007; Owren, Seyfarth, & Cheney, 1997; Pfefferle & Fischer, 2006; Rendall, 2003; Rendall, Kollias, Ney, & Lloyd, 2005). Recently, Fitch and colleagues (Fitch et al., 2016) found that rhesus macaque vocal tract configurations were highly diverse during natural behaviors (i.e., vocalizing, facial displays, and feeding). They concluded that macaques have the capacity to vocalize in a large formant space but lack the neural mechanisms for doing so. In addition, a recent study looking across the vocal repertoire of baboons found that they produce a broad range of sounds indicating extensive articulatory control (Boë et al., 2017). Together, these findings suggest that non-human primates’ ability to modulate the spectral properties of vocalizations may not be as constrained as predicted by the peripheral hypothesis (Fitch et al., 2016; Owren et al., 1997). Still lacking is a comparative understanding of the limits of non-human primates in their ability to modulate f0 and formant profiles, particularly within single vocalizations (Pisanski, Cartei, McGettigan, Raine, & Reby, 2016).

1.2 Temporal features of vocal production in humans and other primates

The human ability to modulate sound extends beyond the spectral domain to the temporal domain. Among the “unique” temporal features of human speech are the ~5 Hz (3-8 Hz range) rhythm in the production rate of small meaningful units like syllables and phonemes bounded by consonants; this speaking rhythm (i.e., 3-8 syllables per second) is facilitated by the controlled movement of facial articulators (e.g., tongue and lips) and breathing (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009; T. H. Crystal & House, 1982; Ghazanfar & Rendall, 2008; Greenberg, Carvey, Hitchcock, & Chang, 2003; MacLarnon & Hewitt, 1999; Malécot, Johnston, & Kizziar, 1972). The 3-8 Hz rhythm is thought to be a universal characteristic of human speech and exists in all languages studied to date, including British English, American English and French (Chandrasekaran et al., 2009). Disrupting the natural speaking rhythm reduces intelligibility (Drullman, 1994; Elliott & Theunissen, 2009; Saberi & Perrott, 1999; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995; Smith, Delgutte, & Oxenham, 2002). This reduction occurs, in part, because the human auditory cortex appears designed to entrain to a speech rhythm in the approximate range of 3-8 Hz (Gross et al., 2013; Peelle & Davis, 2012; Schroeder, Lakatos, Kajikawa, Partan, & Puce, 2008).

Furthermore, the variation in human speech rhythm abides by Menzerath’s law, which states “the greater the whole, the smaller its constituents” (Altmann, 1980; Köhler, 2012; Malécot et al., 1972; Menzerath, 1954). Originally, this law was used to characterize structural properties of written text (Altmann, 1980; Teupenhayn & Altmann, 1984), and it has since been applied to a wide range of complex systems including music and genomic structures (Boroda & Altmann, 1991; Ferrer-i-Cancho & Forns, 2010). Menzerath's law applies to spoken language in that syllable duration decreases (and syllable rate increases) with increasing phrase or utterance lengths (T. H. Crystal & House, 1990; Grégoire, 1899; Lindblom, 1968; Malécot et al., 1972; Nakatani, O'Connor, & Aston, 1981; Quené, 2008; Schwab & Avanzi, 2015). Better understanding of the levels at which Menzerath's law operates will help us gain insight into the role of self-organization (Köhler, 1987) and compression of information (Cramer, 2005; Köhler, 2012) in the evolution of complex vocal systems.

It has been presumed that humans can vocalize at rates 10 times faster than any non-human primate (Lieberman, Laitman, Reidenberg, & Gannon, 1992), and this was thought to be because non-human primates are highly constrained by their breathing and facial articulator abilities (Ghazanfar & Rendall, 2008; Lieberman, 1968; MacLarnon & Hewitt, 1999). Current research challenges these presumptions, however, and a model – the bimodal speech rhythm hypothesis – proposes that rhythmic facial expressions, like lip-smacking, characterized our ancestral primates and set the stage for the fast paced vocalizations that would later become speech (Ghazanfar, 2013; Ghazanfar & Takahashi, 2014). Data from studies on monkey lip-smacking supports this hypothesis by showing that, like humans, rhesus macaques (*Macaca mulata*) move their mouths in a 3-8 Hz rhythm, and this range is preferred by observer monkeys over faster or slower rhythms (Ghazanfar, Morrill, & Kayser, 2013; Ghazanfar, Takahashi, Mathur, & Fitch, 2012). Since then, data on the 'faux-speech' calls of Sumatran orangutans (*Pongo abelii*) and the 'wobble' calls of geladas (*Theropithecus gelada*) show that non-human primates can even go one step further; both species move their lips in a ~3-8 Hz rhythm while vocalizing (Bergman, 2013; Lameira et al., 2015, 2016, 2014). These findings extend the biomodal speech rhythm hypothesis by demonstrating that the coupling of voice to rhythmic facial expressions may not be as complex an evolutionary process as previously thought (Ghazanfar & Takahashi, 2014).

We still lack enough data to build an understanding of the processes shaping rhythm variability in primate vocalizations. In other words, what determines when a primate vocalization will be produced at 3 Hz compared to 8 Hz rhythm? The data on gelada 'wobbles', for example, were limited to wobbles produced on an inhale. This is significant because it is known from human research that vocal tract anatomy operates differently during exhaled and inhaled speech, which results in slower voice-onset times in exhaled speech (Moerman et al., 2016; Ng, Chen, Wong, & Xue, 2011; Vanhecke et al., 2016). As such, a gelada wobble produced on an exhale may have a rhythm that overlaps more or even less with the 3-8 Hz of human speech compared to inhaled wobbles (which range from 5.0-7.5 Hz according to Bergman (2013)). The numbers of wobbles cycles might also make a difference in dictating rhythm. Similar to how Menzerath's law applies to speaking rates in human speech, geladas produce sequences of calls at rates that increase as the sequence size gets larger (Gustison et al., 2016). Yet, even the longest of these call sequences (a 26-call sequence) are produced at the relatively low rate of 3.018 calls per second. It remains to be seen whether the lip movement rhythm of single vocalizations, notably wobbles, gets faster as wobble size increases. Data on how production mode (inhale and exhale) and wobble size (number of lip smacks) influence rhythm will help refine our application of the bimodal rhythm speech hypothesis to human speech evolution.

1.3 A gelada case study

Our goal is to test the boundaries of what makes human speech unique by further exploring the spectro-temporal properties of gelada vocalizations. Geladas, a monkey endemic to the Ethiopian highlands, have been a prime example of a vocally complex non-human primate since studies on them began in the 1970s (Bergman, 2013; Gustison et al., 2012; Richman, 1976, 1987). Early research used case studies to illustrate that gelada calls have distinct vowel qualities, which implies that geladas control the resonance chambers of their vocal tract (Richman 1976). Recent studies show that geladas and baboons share a homologous call type (i.e., exhaled grunts) and that gelada males produce phylogenetically “derived” calls which have acoustic properties (e.g., long duration, larger F1 bandwidth, higher F1 coefficient of variance) that make them more salient than exhaled grunts (Gustison & Bergman, 2016; Gustison et al., 2012). These derived calls are produced almost exclusively by males (Gustison et al., 2012). Research also shows that geladas use quick changes in pitch and consonantal onsets to produce rhythmic units of sound (Richman, 1987). One of the derived gelada calls, the ‘wobble’, is a form of vocalized lip-smacking that has a speech-like rhythm ranging from 6 to 9 Hz (Bergman, 2013). This call is made almost exclusively by males and is the only call with a speech-like rhythm. Due to previous small sample sizes, however, it is unclear whether the periodicity of gelada wobbles follows the temporal organization pattern predicted by Menzerath’s law, although this law has emerged in larger units of gelada communication like call sequences (Gustison et al., 2016).

Here, we investigate the modulation of spectro-temporal properties in discrete gelada vocalizations. First, we measure the formant space areas and contours of gelada calls to better understand their vowel-like qualities. We do this by comparing the F1-F2, F1-F3 and F1-F4 formant space of male and female gelada exhaled grunts to chacma baboons and human vowel space area, and by testing the degree to which geladas vary their f0 and F1-F4 contours relative to baboons. Then, we compare the formant profiles and modulation of the five derived male gelada call types (inhaled grunts, exhaled moans, inhaled moans, wobbles and yawns) to exhaled grunts. These calls are used in the similar context of non-competitive social interactions. Second, we test for evidence of Menzerath's law in the temporal organization of wobble calls to better understand the levels at which this universal principle operates. We do this by associating the cycle duration (i.e., time from mouth closed to mouth closed) with the number of cycles in each wobble. A negative association would support Menzerath's law. We also test whether cycle duration depends on the production mode of wobbles (inhaled vs. exhaled). Finally, we put the findings on gelada spectro-temporal modulation into context by integrating our conclusions with the neural hypothesis and bimodal speech rhythm hypothesis.

2. Methods

2.1 Study sites and animals

Data for this study come from three different bands in one community of wild geladas (about 1200 individuals) living in the Sankaber area of the Simien Mountains National Park, Ethiopia (2008–2014) and a single group of chacma baboons (group C) living in the Moremi Game Reserve in the Okavango Delta of Botswana (2001–2002). The gelada units were comprised of one leader male, 0–3 follower males, and 1–11 females and their immature offspring. The chacma baboon group ranged from 82 to 91 individuals, including 9–11 adult males, 29–31 adult females and their immature offspring. All subjects were habituated to humans on foot up to 3-5 m and could be identified by unique body markings (e.g., ear tears and coloration). Research was approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan, the Institutional Animal Care and Use Committee (IACUC) at the University of Pennsylvania, and was carried out in accordance with the laws and approved guidelines of the Ethiopian government.

2.2 Acoustic recordings and processing

We opportunistically recorded vocalizations from 25 adult male and 32 adult female geladas (Feb 2008–Jun 2014) and 9 adult male and 15 adult female chacma baboons (Apr 2001– May 2002) with a Sennheiser ME66 directional microphone connected to a digital stereo recorder (Marantz PMD 660 and 661 Digital Recorder for geladas; Sony VW-D6 Professional Walkman for chacma baboons). Call recordings were chosen for spectro-temporal analyses if they were free of background noise (e.g., wind and other animal calls). A subset of these recordings have been used in other papers for different types of analyses (Bergman, 2013; Gustison et al., 2012, 2016). The call types and contexts of all vocalizations were described at the time of recording. Our analyses focused on chacma baboon and gelada exhaled grunts, as well as gelada male derived calls (inhaled grunts, exhaled moans, inhaled moans, wobbles, and yawns); these are the call types that occurred repeatedly during affiliative interactions and foraging (Figure 3.1). Previous studies show that these call types have a high inter-observer reliability of 96% (Gustison et al., 2012). Because gelada grunts and moans (an elongated version of a grunt) grade into each other somewhat, density plots of call durations (log-transformed) for exhaled and inhaled grunts/moans were used to determine thresholds to distinguish between grunt and moan call types (Gustison et al., 2016). These density plots show bimodal distributions. The threshold (lowest point between the two peaks) between exhaled grunts and moans was 0.768 s and the threshold between inhaled grunts and moans was 0.513 s.

2.3 Spectro-temporal parameters

Extraction of spectro-temporal parameters was performed using Praat© acoustic software (version 6.0.23) on the Macintosh OX operating system (Boersma & Weenink, 2011). First, we made comparisons of fundamental frequency (f0) and formant profiles (F1-F4) across species and gelada male call types. To obtain f0 contours, we used a custom Praat script that applied Praat's autocorrelation algorithm to extract points every 6.25 ms. We set a broad search range of 25-500 Hz for chacma baboon grunts and gelada grunts, moans and wobbles. We set a search range of 25-1000 Hz for gelada yawns since they covered a larger f0 range. All call recordings were set to a 11025 Hz sampling frequency with no filters using Avisoft SAS Pro 5.2 (R. Specht, Berlin, Germany). f0 contours were checked manually for outliers and any artificial pitch jumps. Then, edited f0 contours were used to calculate the mean f0 frequency and f0 bandwidth (maximum f0 frequency minus minimum f0 frequency) for each call.

We identified formant contours (F1-F4) with linear predictive coding (LPC) using the burg method in Praat. We define formants as peaks in the frequency spectrum. LPC estimates the frequency values of formants based on the assumption that the vocal signal is produced by a buzz generated at the glottis or source and is shaped into resonances as it passes through the vocal tract filter (throat, mouth and nose) (Owren & Bernacki, 1998; Taylor et al., 2016). Without anatomical measurements of the vocal tract, however, it is important to note that we cannot say that these identified formants are precise estimates of vocal tract resonances. For further discussion on the definitions of formants and vocal tract resonances, please see Titze et al (2015). Before carrying out LPC analyses, we reduced the sampling frequencies to 11025 Hz (the nyquist frequency) so that the corresponding frequency range would be 5512.5 Hz. The only exception to this were gelada females, for which we used sampling frequencies of 22050 Hz. This sampling frequency was higher because female gelada grunts sometimes had harmonics that ranged beyond 5512.5Hz. We used a 50 Hz high-pass IIF filter to dampen any lingering low-frequency background noise. Sampling frequency conversion and filtering was carried out with Avisoft SASLab Pro. This frequency range was chosen because it was the highest frequency found in the majority of gelada and chacma baboon recordings in our dataset, and it is a standard range used for studies of formants in baboons (Pfefferle & Fischer, 2006; Rendall et al., 2005).

We used a Praat to extract data from LPC spectrum slices every 6.25 ms with a frequency analysis window of 25 ms and coefficient settings ranging from 14-18. We identified the first four formants (F1-F4) per slice by locating the peaks in the LPC spectra across the 0-5512.5 Hz range. Two methods were used to check the goodness of fit of formants predicted by the LPC slices. First, LPC spectra were overlaid on three independently derived FFT spectra (512-point Hanning window with a pre-emphasis filter) at one-, two- and three-quarters of the way through each call (512-point Hanning window). Second, the formant contours were compared with the actual call spectrograms (512-point Hanning window). All recordings used in the following analyses had at least three detectable formant contours. Any outliers (i.e., points outside of the formant median ± 2 SEM) in the formant contours were removed so that there were no artificial frequency jumps. Then, edited F1-F4 contours were used to calculate the mean formant frequencies and bandwidths (maximum formant frequency minus minimum formant frequency) for each call.

Durations of inhaled wobble and exhaled wobble cycles were measured with Praat (Figure 3.2a-b). A cycle corresponds to an opening and closing of the mouth. For an visual example of a wobble, see the supplementary movie published in (Bergman, 2013). To quantify cycle duration, we created ‘intensity tiers’ (Bergman, 2013) from a wobble using a minimum frequency of 100 Hz. Local amplitude minimums in these intensity contours were identified manually (Figure 3.2c). Cycle durations were then calculated as the time between local amplitude minimums (the start of a wobble was counted as a local amplitude minimum).

2.4 Data analysis

Mean f0, F1, F2, F3, and F4 frequencies and bandwidths were determined for each individual per call type. Individuals were retained for data analyses if they had available data for f0 and each of the four formant measurements. Formant dispersion

(D_f), was then calculated as the average distance between adjacent pairs of formant frequencies using the following formula (Fitch, 1997):

$$D_f = \frac{\sum_{i=1}^{N-1} F_{i+1} - F_i}{N-1}, \qquad (1)$$

where D_f is formant dispersion (in Hz), N is the total number of formants measured, and F_i is the frequency (in Hz) of formant i. D_f was then used to estimate vocal tract length (VTL, in cm) using the following formula (Taylor et al., 2016):

$$VTL = \frac{c}{2 * D_f} * 100, \qquad (2)$$

where c is the speed of sound in air (~ 350 m/s in the warm humid air of a mammalian vocal tract).

Formant space was determined for each formant pair (F1-F2, F1-F3, F1-F4) by calculating the ellipse of the 95% confidence region that was formed with the "dataEllipse" function in the R software package "car" (Fox & Weisberg, 2011). These formant space areas were compared across gelada and chacma baboon exhaled grunts and human English vowels. Vowel formant frequencies in men and women were obtained from Table 5 in Hillenbrand et al. (1995). Then, two-way ANOVAs were used to compare f0, F1, F2, F3, and F4 mean frequencies and mean bandwidths across species (chacma baboon and gelada), sex and the interaction between species and sex. Linear Mixed Models (LMM) were used to compare f0, F1, F2, F3, and F4 mean frequencies and mean bandwidths across gelada male call types. In these comparisons, call type (exhaled grunt, inhaled grunt, exhaled grunt, exhaled moan, wobble, yawn) was the fixed effect. We investigated differences between call types using pairwise comparisons based on the LMM output. For the above analyses, we combined inhaled and exhaled wobble recordings into one ‘wobble’ category due to low sample size of high quality wobble recordings.

To investigate the temporal dynamics of wobble calls, we first used Spearman rank correlation tests, a method introduced to examine Menzerath’s law and remain objective about the exact functional dependency between variables (Baixeries, Hernández-Fernández, Forns, & Ferrer-i-Cancho, 2013; Ferrer-i-Cancho et al., 2014; Nikolaou, 2014). In these tests, cycle durations were correlated with the size (number of total cycles) of their associated wobble. If Menzerath’s law applies to wobbles, then there should be a negative association between cycle duration and wobble size. We also tried this analysis for each production mode (exhaled or inhaled) separately because it is not yet known whether the cycle length depends on the way in which a wobble is vocalized. Previous research focused on inhaled wobbles (Bergman, 2013). Although use of Spearman rank correlations can help to avoid potential problems of previous research on Menzerath’s law (Ferrer-i-Cancho et al., 2014), we recognize that its application with our dataset involves pseudo-replication. Therefore, we used a LMM to test whether the variation in cycle duration corresponds to production mode (exhaled or inhaled) or wobble size (number of cycles in a wobble) while controlling for the position of the cycle within a wobble (first, second, third, etc.), the wobble recording, and caller identity. In this LMM, cycle duration was the dependent variable, wobble size and production mode (exhaled or inhaled) were fixed effects, and cycle position, wobble recording, and caller identity were random effects.

We ran the LMMs using the function “lmer” of the R package “lme4” (Bates, Maechler, & Bolker, 2012). The “lmerTest” package was implemented to determine the significance of the LMM coefficients (Kuznetsova, Brockhoff, & Christensen, 2013). We used the “difflsmeans” function in the lmerTest package to make pairwise comparisons between call types. Unless noted otherwise, all calculations and statistical tests were carried out in R 3.3.0 (R Development Core Team, 2016). Critical values were set at alpha = 0.05, and all tests were two-tailed.

3. Results

3.1 Comparison of spectral properties across species

The following analyses are based on the f0 and formant measurements of 280 exhaled grunts from 73 individuals (Table 3.1). The formant space of adult gelada exhaled grunts covered higher frequencies compared to chacma baboons and overlapped to a large degree with human vowel space (Table 3.1; Figure 3.3a-c). ANOVAs revealed that mean formant frequencies differed across species and sex. Geladas had higher mean f0 ($F_{1,69} = 528.98$, $P < 0.0001$), F1 ($F_{1,69} = 227.98$, $P < 0.0001$), F2 ($F_{1,69} = 69.37$, $P < 0.0001$), F3 ($F_{1,69} = 104.49$, $P < 0.0001$), and F4($F_{1,69} = 210.71$, $P < 0.0001$) frequencies compared to chacma baboons. Females had higher mean f0 ($F_{1,69} = 97.52$, $P < 0.0001$), F1 ($F_{1,69} = 30.43$, $P < 0.0001$), F2 ($F_{1,69} = 36.78$, $P < 0.0001$), F3 ($F_{1,69} = 82.55$, $P < 0.0001$), and F4 ($F_{1,69} = 116.78$, $P < 0.0001$) frequencies compared to males. No species-sex interactions were found for mean f0 ($F_{1,69} = 3.37$, $P = 0.0709$), F1 ($F_{1,69} = 0.24$, $P = 0.6233$), F2 ($F_{1,69} = 2.12$, $P = 0.1503$) or F3 ($F_{1,69} = 2.51$, $P = 0.1176$) frequencies. There was an interaction for mean F4 in that gelada and chacma females were more different from each other than were males ($F_{1,69} = 17.50$, $P = 0.0001$).

ANOVAs also showed that f0 and formant modulation, as measured by bandwidth for each call per formant, differed across species and sex (Figure 3.3d-h). Geladas had larger mean bandwidths of f0 ($F_{1,69} = 36.23$, $P < 0.0001$), F1 ($F_{1,69} = 79.12$, $P < 0.0001$), F2 ($F_{1,69} = 76.62$, $P < 0.0001$), F3 ($F_{1,69} = 46.84$, $P < 0.0001$), and F4 ($F_{1,69} = 42.49$, $P < 0.0001$) compared to chacma baboons. Compared to males, females had larger mean bandwidths for f0 ($F_{1,69} = 10.50$, $P = 0.0018$) and F1 ($F_{1,69} = 28.80$, $P < 0.0001$), but there were no clear sex differences for F2 ($F_{1,69} = 3.33$, $P = 0.0725$), F3 ($F_{1,69} = 1.82$, $P = 0.182$) or F4($F_{1,69} = 0.06$, $P = 0.808$). No species-sex interactions were found for bandwidths of f0 ($F_{1,69} = 3.17$, $P = 0.0793$), F2 ($F_{1,69} = 0.66$, $P = 0.4199$), F3 ($F_{1,69} = 0.47$, $P = 0.4949$), or F4 ($F_{1,69} = 3.38$, $P = 0.0702$). There was an interaction for F1 bandwidth in that gelada and chacma females were more different from each other than were males ($F_{1,69} = 12.69$, $P = 0.0007$).

3.2 Comparison of spectral properties across gelada calls

The following analyses are based on the f0 and formant measurements of exhaled grunts (n = 107 calls), inhaled grunts (n = 66 calls), exhaled moans (n = 14 calls), inhaled moans (n = 32 calls), wobbles (n = 15 calls), and yawns (n = 19 calls) made by 25 gelada males (Table 3.1-3.2). LMMs revealed that exhaled grunts and derived calls had different formant profiles (Table 3.3, Figure 3.4a,c,e,g,i) and levels of within-call modulation (Table 3.4, Figure 3.4c,d,f,h,j). There was no evidence that mean frequencies and bandwidths differed between exhaled and inhaled grunts. Exhaled moans had higher mean F2 frequencies than exhaled grunts. Inhaled moans had lower mean f0 and F1 frequencies than exhaled grunts. Wobbles had lower mean F2-F4 frequencies and larger F1-F3 bandwidths than exhaled grunts. Yawns had lower mean f0 and F1-4 frequencies and larger f0 and F1-4 bandwidths than exhaled grunts.

3.3 Temporal properties of gelada wobble calls

We analyzed 28 inhaled and 25 exhaled wobble recordings (composed of a total of 239 cycles) recorded from 25 study males (1–14 wobble recordings per male). The wobbles ranged between 2—17 cycles and a total duration of 0.376—2.763 seconds. Cycle duration averaged 0.204 ± 0.006 seconds (4.912 Hz) and ranged from 0.555–0.056 seconds (1.802–17.857 Hz) (Figure 3.5). Inhaled wobbles ranged between 2—7 cycles (mean duration of 0.180 ± 0.007 seconds (5.569 Hz) and range of 0.555–0.080 seconds (1.802—12.500 Hz). Exhaled wobbles ranged between 2—17 cycles (mean duration of 0.221 ± 0.008 seconds (4.522 Hz) and range of 0.544—0.056 seconds (1.838—17.857 Hz).

We first used Spearman rank correlations to test how the durations of wobble cycles were related to the size (number of cycles) of their corresponding wobbles. We found a negative correlation between cycle duration and wobble size (rho = -0.156, $P = 0.0156$; Figure 3.5a). This negative association was also characteristic of exhaled wobbles only (rho = -0.330, $P < 0.0001$) and inhaled wobbles only (rho = -0.410, $P < 0.0001$). We then used a LMM to investigate how variability in cycle duration was associated with both wobble size and production mode (exhaled and inhaled) while including cycle position, caller identity, and wobble recording as random effects. Corroborating the findings of the Spearman correlation, the LMM showed a negative association between cycle duration and wobble size (Estimate ± SEM = -0.008 ± 0.003, *t* = 2.731, *P* = 0.0138; Figure 3.6A). Cycles were shorter in inhaled compared to exhaled wobbles (Estimate ± SEM = -0.065 ± 0.016, *t* = 3.951, *P* = 0.003; Figure 3.5b). To test whether the slopes between wobble size and cycle duration differed based on wobble type, we ran a separate model to investigate whether there was an interaction between wobble type and size. We found no evidence that this was the case (Estimate ± SEM = -0.010 ± 0.009, *t* = 1.112, *P* = 0.2721).

4. Discussion

Gelada vocalizations exhibited spectro-temporal dynamics that had several similarities to the modulatory capacity of human speech as compared to the vocalizations of their baboon relatives. We found that gelada exhaled grunts covered a formant space area that overlapped more with human vowel space area (VSA) than did chacma baboon grunts. The f0 and first four formants were more modulated in gelada grunts than in chacma baboon grunts. Also, the exhaled grunts and derived calls showed diverse formant profiles. Wobbles and yawns had particularly high levels of f0 and formant modulation. We also found that wobbles, vocalized lip-smacking with a speech-like rhythm, has a rhythm that overlaps more with human speech than previously thought. Wobble rhythm depends on the production mode (inhaled vs. exhaled) of the call and the size (number of cycles in a wobble). Following Menzerath's law, wobble cycle duration was shorter if the corresponding wobble size was larger. Together, these findings suggest that geladas have a significant capacity to modulate the spectro-temporal dynamics of their vocalizations, which in turn, has important implications for proposed hypotheses on the evolution of a speech-like ability in primates.

Previous research on the vocal capabilities of our non-human primate relatives have supported one of two hypotheses, the “peripheral” hypothesis or the “neural” hypothesis. The neural hypothesis advocates that non-human primates have the anatomical capacity but lack the neural mechanisms to flexibly control how they modulate their sounds (Fitch et al., 2016; Hockett, 1960). The present study falls in line with several other recent studies on apes and old world monkeys supporting the neural hypothesis (Boë et al., 2017; Fitch et al., 2016; Lameira et al., 2016, 2014). Our findings support the neural hypothesis in three ways. First, our cross-species comparison suggests that geladas have a greater ability to control the spectral properties of grunt calls than do baboons, whose grunts were already thought to cover a substantial degree of formant space (Boë et al., 2017; Owren et al., 1997; Rendall et al., 2005). The formant space of gelada grunts overlapped more with human VSA than baboon grunts. Specifically, gelada male formant space fell in the VSA shared by men and women, and female geladas covered most of women’s VSA. Moreover, gelada males and females varied the formant (and f0) frequencies within individual grunts to a level that exceeded the relatively unmodulated grunts made by chacma baboons. One potential caveat of the present study is that formants were measured by finding peaks in spectrum rather than vocal tract resonances, and so, there are limits to comparisons that can be made with human VSA. Modeling research on the vocal anatomy of modern humans, Neanderthals and rhesus macaques suggests that a combination of physical traits can result in larger formant space: a lower larynx, increased pharynx size, and enhanced gestures of the tongue body (lips and jaw) (Boë, Heim, Honda, & Maeda, 2002; Lieberman & Crelin, 1971; Lieberman et al., 1969). Recent studies in macaques and baboons suggest that human-like anatomy is not necessary for producing calls that cover a wide formant space (Boë et al., 2017; Fitch et al., 2016). However, without morphological data on gelada vocal tract anatomy, we do not know what, if any, anatomical adaptations underlie their diverse vocalizations. Regardless, the current findings still suggest that geladas have a greater tendency to vary the spectral properties of their calls than do chacma baboons.

The second way in which our data support the neural hypothesis is through comparisons of the homologous and derived call types produced by gelada males. Exhaled grunts (the homologous call type shared with baboons) had a similar degree of f0 and formant modulation as inhaled grunts and inhaled/exhaled moans, but they were less modulated than wobbles and yawns. Additionally, the exhaled grunts and most of the derived calls had unique formant profiles. The only call types that had similar formant profiles were exhaled and inhaled grunts. The diversity in formant profiles across these six call types is important because these are the calls used to form call sequences (Gustison et al., 2016). Thus, a given vocal sequence will contain enhanced spectral modulation if it combines derived call types with exhaled/inhaled grunts. Presumably, geladas must exhibit some control over their supralaryngeal vocal tract so they can flexibly filter sound across a single call or call sequence. Furthermore, the variability that we describe is almost certainly underestimating the capabilities of geladas as we focus on only those calls used in the same contexts (affiliative social interactions) as grunts. Including more acoustically divergent calls such as alarm calls, screams, and display calls, would likely expand the formant space covered by geladas as it does in baboons (Boë et al., 2017).

The third way in which our data support the neural hypothesis is through the speech-like rhythmic capacity of gelada male wobble calls. We replicated previous findings showing that gelada males move their facial articulators (e.g., mouth) during wobbles in a similar rhythm to human speech. By using an larger dataset than previous research (Bergman, 2013), we show that the rhythm of this vocalized lip-smacking overlaps more with human speech rhythm than previously thought. In the present study, wobble rhythm ranged from 2 to 15 Hz, which overlaps entirely with the 3-8 Hz rhythm of speech (Chandrasekaran et al., 2009). Moreover, this wobble rhythm depends on the production mode, with exhaled wobbles having longer cycles on average than inhaled wobbles. Furthermore, we found evidence that the variation in wobble rhythm abides by Menzerath's law (Grégoire, 1899; Quené, 2008) in that there is a negative association between wobble cycle length and wobble size. In other words, wobble tempo got faster as wobble size got larger. Menzerath's law also applied to inhaled and exhaled wobbles separately, although it should be noted that it is unclear the degree to which this pattern in exhaled wobbles is driven by the cycle durations of an unusually large 17-cycle wobble (Figure 3.5). The presence of Menzerath's law in gelada vocalizations also has been discovered at the level of call sequences in that calls from larger sequences had shorter durations than calls from smaller sequences (Gustison et al., 2016). Our current findings therefore demonstrate that Menzerath's law, and by extension the compression of vocal signals, operates at multiple levels of gelada communication. Additionally, these speech-like patterns of rhythm modulation in gelada calls support the neural hypothesis and expand on the bimodal speech rhythm hypothesis. The bimodal speech rhythm hypothesis proposes that primate rhythmic facial expressions set the stage for fast paced vocalizations (Ghazanfar, 2013). The current findings expand on this hypothesis by showing that geladas not only have the capacity to couple vocalization to rhythmic facial expression, but that “linguistic” laws of compression predict the variability in this rhythm. Thus, the coupling of voice to facial movement is not an adaptation specific to human speech.

Our findings have implications for research on the selective pressures driving derived vocal traits in humans and geladas. In humans, listeners have a difficult time understanding speech that has smaller vowel space, reduced articulation, and abnormal speech tempos (Bradlow, Torretta, & Pisoni, 1996; Neel, 2008). These speech deficits are characteristic of people with dysarthria, Parkinson’s, depression, and PTSD, for example (Levy et al., 2015; Scherer, Lucas, Gratch, Rizzo, & Morency, 2016; Skodda, Grönheit, & Schlegel, 2012; Volkmann, Hefter, Lange, & Freund, 1992; Whitfield & Goberman, 2014). Thus, expanded VSA and moderately paced syllable rates are crucial for the effective transfer of complex information. It does not appear that the expanded formant space and speech-like rhythm in the gelada male vocal system has led to enhanced referential information transfer. Rather, social functions appear more important. A playback study showed that gelada females appear to pay more attention to male utterances that included derived call types with the greatest formant modulation – exhaled moans, wobbles and yawns (Gustison & Bergman, 2016). Considering that there is a sex-bias in the production of these call types, it is likely that sexual selection played a role in driving the morphological and/or behavioral traits needed to produce complex sounds (Gustison et al., 2012). These findings align with human studies showing that men tend to use a larger and more unique vocabulary around women, and women tend to prefer men with a larger vocabulary (Prokosch, Coss, Scheib, & Blozis, 2009; Rosenberg & Tunney, 2008).

There is notable debate over which traits of speech are unique to humans (Boë et al., 2002; Fitch, 2010; Ghazanfar & Rendall, 2008; MacLarnon & Hewitt, 2004). Our findings on the spectro-temporal properties of gelada calls add to this ongoing conversation by building on hypotheses about the evolution of complex vocal ability. By exhibiting a high degree of modulation in the spectrum-based and temporal-based features of their calls, as compared to baboons, geladas challenge the traditional hypothesis that non-human primates lack the physical capacity to produce complex speech-like sounds. Instead, our findings support the alternative hypothesis that primates have a capacity to produce complex sounds but lack the appropriate neural mechanisms to fully control this capacity. The distinction between humans and other primates on the basis of vocal ability may be smaller than previously assumed. Future work will be essential in identifying selective pressures that may have contributed to the evolution of these modulatory skills in geladas and other non-human primates.

Acknowledgements

MLG was supported by grants from the National Geographic Society (Gr# 9122-12, W304-14), Leakey Foundation, University of Michigan, and Petridish (primary supporter: John Allen). TJB and the gelada project were supported by the Wildlife Conservation Society (SSF Gr# 67250), the National Geographic Society (Gr# 8100-06), the Leakey

Foundation, the National Science Foundation (Gr# BCS-0715179, BCS-0962118, IOS-1255974), and the University of Michigan. Research in Botswana was supported by NIH grant MH62249, an NRSA fellowship to TJB, the Leakey Foundation and the University of Pennsylvania. We are grateful to the Ethiopian Wildlife Conservation Authority and the wardens and staff of the Simien Mountains National Park for granting us permission to conduct this research in Ethiopia. We thank the Office of the President and the Department of Wildlife and National Parks of the Republic of Botswana for permission to conduct research in the Moremi Reserve. We thank Dorothy L. Cheney and Robert M. Seyfarth and for permission to use recordings of chacma baboon grunts and Aliza le Roux for use of gelada wobble and grunt recordings. We also thank Schuyler Cohen for help constructing the customized Praat scripts.

References

Aich, H., Moos-Heilen, R., & Zimmermann, E. (1990). Vocalizations of adult gelada baboons (*Theropithecus gelada*): Acoustic structure and behavioural context. *Folia Primatologica*, *55*, 109–132.

Altmann, G. (1980). Prolegomena to Menzerath's law. *Glottometrika*, *2*, 1–10.

Andrew, R. J. (1976). Use of formants in the grunts of baboons and other nonhuman primates. *Annals of the New York Academy of Sciences*, *280*, 673–93.

Baixeries, J., Hernández-Fernández, A., Forns, N., & Ferrer-i-Cancho, R. (2013). The parameters of the Menzerath-Altmann Law in genomes. *Journal of Quantitative Linguistics*, *20*, 94–104.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. Retrieved from http://cran.r-project.org/ package=lme4

Benítez, M. E., le Roux, A., Fischer, J., Beehner, J. C., & Bergman, T. J. (2016). Acoustic and temporal variation in gelada (Theropithecus gelada) loud calls advertise male quality. *International Journal of Primatology*. http://doi.org/10.1007/s10764-016-9922-0

Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Current Biology : CB*, *23*, R268-9.

Boë, L.-J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., … Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (Papio papio) suggests pre-hominin speech precursors. *PloS One*, *12*, e0169321.

Boë, L.-J., Heim, J.-L., Honda, K., & Maeda, S. (2002). The potential Neandertal vowel space was as large as that of modern humans. *Journal of Phonetics*, *30*, 465–484.

Boersma, P., & Weenink, D. (2011). Praat: doing phonetics by computer. Retrieved from http://www.fon.hum.uva.nl/praat/

Boroda, M. G., & Altmann, G. (1991). Menzerath's law in musical texts. In M. G. Boroda (Ed.), *Musikometrika 3* (pp. 1–13). Bochum: Brockmeyer.

Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication*, *6393*, 255–272.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A.

(2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, *5*, e1000436.

Chiba, T., & Kajiyama, M. (1941). *The vowel: Its nature and structure*. Tokyo: Tokyo-Kaiseikan.

Cramer, I. (2005). The parameters of the Altmann-Menzerath law. *Journal of Quantitative Linguistics*, *12*, 41–52.

Crystal, D. (2003). *The Cambridge Encyclopedia of Language* (2nd ed.). Cambridge: Cambridge University Press.

Crystal, T. H., & House, A. S. (1982). Segmental durations in connected speech signals: preliminary results. *Journal of the Acoustic Society of America*, *72*, 705–716.

Crystal, T. H., & House, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *Journal of the Acoustic Society of America*, *88*, 101–112.

Drullman, R. (1994). Effect of temporal envelope smearing on speech reception. *The Journal of the Acoustical Society of America*, *95*, 1053.

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Computational Biology*, *5*, e1000302.

Fant, G. (1960). Acoustic theory of speech production. The Hague, The Netherlands: Mouton & Co.

Ferrer-i-Cancho, R., & Forns, N. (2010). The self-organization of genomes. *Complexity*, *15*, 34–36.

Ferrer-i-Cancho, R., Hernández-fernández, A., Baixeries, J., Debowski, L., Macutek, J., & Hernàndez-Feràndez, A. (2014). When is Menzerath-Altmann law mathematically trivial? A new approach. *Statistical Applications in Genetics and Molecular Biology*, *13*, 633–644.

Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, *102*, 1213–1222.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, *4*, 258–267.

Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University

Press.

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, *2*, e1600723.

Fitch, W. T., & Hauser, M. D. (1995). Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on “Honest” advertisement. *American Journal of Primatology*, *37*, 191–219.

Fitch, W. T., & Suthers, R. A. (2016). Vertebrate vocal production: An introductory overview. In *Vertebrate Sound Production and Acoustic Communication* (pp. 1–18). Springer International Publishing. http://doi.org/10.1007/978-3-319-27721-9

Fox, J., & Weisberg, S. (2011). *An {R} companion to applied regression, 2nd edition*. Thousand Oaks, CA: Sage.

Ghazanfar, A. A. (2013). Multisensory vocal communication in primates and the evolution of rhythmic speech. *Behavioral Ecology and Sociobiology*, *67*, 1441–1448.

Ghazanfar, A. A., Morrill, R. J., & Kayser, C. (2013). Monkeys are perceptually tuned to facial expressions that exhibit a theta-like speech rhythm. *Proceedings of the National Academy of Sciences*, *110*, 1959–63.

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Current Biology*, *18*, 457–460.

Ghazanfar, A. A., & Takahashi, D. Y. (2014). The evolution of speech: Vision, rhythm, cooperation. *Trends in Cognitive Sciences*, *18*, 543–553.

Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., & Fitch, W. T. (2012). Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Current Biology*, *22*, 1176–1182.

Ghazanfar, A. A., Turesson, H. K., Maier, J. X., van Dinther, R., Patterson, R. D., & Logothetis, N. K. (2007). Vocal-tract resonances as indexical cues in rhesus monkeys. *Current Biology*, *17*, 425–430.

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech - A syllable-centric perspective. *Journal of Phonetics*, *31*,

Grégoire, A. (1899). Variation de la durée de la syllabe en français. *Le Parole*, *1*, 161–176.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., & Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biology*, *11*, e1001752.

Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Scientific Reports*, *6*, 19680.

Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (Theropithecus gelada) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1847–59.

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*, E2750–E2758.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*, 3099–3111.

Hockett, C. (1960). The origin of speech. *Scientific American*, *203*, 88–96.

Köhler, R. (1987). System theoretical linguistics. *Theoretical Linguistics*, *14*, 241–258.

Köhler, R. (2012). Quantitative syntax analysis. Berlin: De Gruyter Mouton.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/package=lmerTest

Lameira, A. R., Hardus, M. E., Bartlett, A. M., Shumaker, R. W., Wich, S. A., & Menken, S. B. J. (2015). Speech-like rhythm in a voiced and voiceless orangutan call. *PLoS ONE*, *10*, e116136.

Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A., & Shumaker, R. W. (2016). Vocal fold control beyond the species-specific repertoire in an orang-utan. *Scientific Reports*, *6*, 30315.

Lameira, A. R., Maddieson, I., & Zuberbühler, K. (2014). Primate feedstock for the evolution of consonants. *Trends in Cognitive Sciences*, *18*, 60–62.

Levy, E. S., Leone, D., Moya-Gale, G., Hsu, S.-C., Chen, W., & Ramig, L. O. (2015). Vowel intelligibility in children with and without dysarthria: An exploratory study.

Communication Disorders Quarterly, *37*, 171–179.

Lieberman, P. H. (1968). Primate vocalizations and human linguistic ability. *Journal of the Acoustic Society of America*, *44*, 1574–1584.

Lieberman, P. H., & Crelin, E. S. (1971). On the speech of Neanderthal man. *Linguistic Inquiry*, *2*, 203–222.

Lieberman, P. H., Crelin, E. S., & Klatt, D. H. (1972). Phonetic ability and related anatomy of the newborn and adult human , Neanderthal man, and the chimpanzee. *American Anthropologist*, *74*, 287–307.

Lieberman, P. H., Klatt, D. H., & Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, *164*, 1185–1187.

Lieberman, P. H., Laitman, J. T., Reidenberg, J. S., & Gannon, P. J. (1992). The anatomy, physiology, acoustics and perception of speech: Essential elements in analysis of human speech. *Journal of Human Evolution*, *23*, 447–467.

Lindblom, B. (1968). Temporal organization of syllable production. *Transmission Laboratory Quarterly Progress and Status Report (Royal Institute of Technology, Stockholm)*, *9*, 1–5.

MacLarnon, A. M., & Hewitt, G. P. (1999). The evolution of human speech: the role of enhanced breathing control. *American Journal of Physical Anthropology*, *109*, 341–63

MacLarnon, A. M., & Hewitt, G. P. (2004). Increased breathing control: Another factor in the evolution of human language. *Evolutionary Anthropology*, *13*, 181–197.

Maddieson, I. (1984). *Patterns of sounds (Cambridge Studies in Speech Science and Communication)*. Cambridge, UK: Cambridge University Press.

Malécot, A., Johnston, R., & Kizziar, P.-A. (1972). Syllabic rate and utterance length in French. *Phonetica*, *26*, 235–251.

Menzerath, P. (1954). *Die Architektonik des deutschen Wortschatzes*. Dümmler: Bonn.

Moerman, M., Vanhecke, F., Van Assche, L., Vercruysse, J., Daemers, K., & Leman, M. (2016). Vocal Tract Morphology in Inhaling Singing: An MRI-Based Study. *Journal of Voice*, *30*, 466–471.

Nakatani, L. H., O'Connor, K. D., & Aston, C. H. (1981). Prosodic aspects of American

English speech rhythm. *Phonetica*, *38*, 84–106.

Neel, A. T. (2008). Vowel space characteristics and vowel identification accuracy. *Journal of Speech, Language, and Hearing Research*, *51*, 574–585.

Ng, M. L., Chen, Y., Wong, S., & Xue, S. (2011). Interarticulator timing control during inspiratory phonation. *Journal of Voice*, *25*, 319–325.

Nikolaou, C. (2014). Menzerath-Altmann law in mammalian exons reflects the dynamics of gene structure evolution. *Computational Biology and Chemistry*, *53*, 134–143.

Owren, M. J., & Bernacki, R. H. (1998). Applying linear predictive coding (LPC) to frequency-spectrum analysis of animal acoustic signals. In S. L. Hopp, M. J. Owren, & C. S. Evans (Eds.), *Animal acoustic communication: sound analysis and research methods* (pp. 129–162). Berlin Heidelberg: Springer-Verlag.

Owren, M. J., Seyfarth, R. M., & Cheney, D. L. (1997). The acoustic features of vowel-like grunt calls in chacma baboons (Papio cyncephalus ursinus): implications for production processes and functions. *The Journal of the Acoustical Society of America*, *101*, 2951–63.

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology*, *3*, 320.

Perlman, M., & Clark, N. (2015). Learned vocal and breathing behavior in an enculturated gorilla. *Animal Cognition*, *18*, 1165–1179.

Pfefferle, D., & Fischer, J. (2006). Sounds and size: identification of acoustic variables that reflect body size in hamadryas baboons, Papio hamadryas. *Animal Behaviour*, *72*, 43–51.

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice modulation: A window into the origins of human vocal control? *Trends in Cognitive Sciences*, *20*, 304–318.

Prokosch, M. D., Coss, R. G., Scheib, J. E., & Blozis, S. A. (2009). Intelligence and mate choice: intelligent men are always appealing. *Evolution and Human Behavior*, *30*, 11–20.

Quené, H. (2008). Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo. *The Journal of the Acoustical Society of America*, *123*, 1104–1113.

R Development Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Raphael, L. J., Borden, G. J., & Harris, K. S. (2007). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Philadelphia: Lippincott Williams & Wilkins.

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*, 3390–3402.

Rendall, D., Kollias, S., Ney, C., & Lloyd, P. (2005). Pitch (F[0]) 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*,

Richman, B. (1976). Some vocal distinctive features used by gelada monkeys. *The Journal of the Acoustical Society of America*, *60*, 718–24.

Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates*, *28*, 199–223.

Rosenberg, J., & Tunney, R. J. (2008). Human vocabulary sse as display. *Evolutionary Psychology*, *6*, 538–549.

Saberi, K., & Perrott, D. R. (1999). Cognitive restoration of reversed speech. *Nature*, *398*, 760.

Scherer, S., Lucas, G. M., Gratch, J., Rizzo, A., & Morency, L. P. (2016). Self-reported symptoms of depression and PTSD are associated with reduced vowel space in screening interviews. *IEEE Transactions on Affective Computing*, *7*, 59–73.

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*(3), 106–113.

Schwab, S., & Avanzi, M. (2015). Regional variation and articulation rate in French. *Journal of Phonetics*, *48*, 96–105.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarilty temporal cues. *Science*, *270*, 303–304.

Skodda, S., Grönheit, W., & Schlegel, U. (2012). Impairment of vowel articulation as a possible marker of disease progression in parkinson’s disease. *PLoS ONE*, *7*, e32132.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*, 87–90.

Taylor, A. M., Charlton, B. D., & Reby, D. (2016). Vertebrate sound production and acoustic communication. In R. A. Suthers, W. T. Fitch, R. R. Fay, & A. N. Popper (Eds.), *Vertebrate Sound Production and Acoustic Communication* (pp. 83–117). Springer International Publishing.

Taylor, A. M., & Reby, D. (2010). The contribution of source-filter theory to mammal vocal communication research. *Journal of Zoology*, *280*, 221–236.

Teupenhayn, R., & Altmann, G. (1984). Clause length and Menzerath’s law. *Glottometrika*, *6*, 127–138.

Titze, I. R., Baken, R. J., Bozeman, K. W., Granqvist, S., Henrich, N., Herbst, C. T., … Wolfe, J. (2015). Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. *The Journal of the Acoustical Society of America*, *137*, 3005–3007.

Vanhecke, F., Lebacq, J., Moerman, M., Manfredi, C., Raes, G. W., & DeJonckere, P. H. (2016). Physiology and acoustics of inspiratory phonation. *Journal of Voice*, *30*, 769.e9-769.e18.

Volkmann, J., Hefter, H., Lange, H. W., & Freund, H. J. (1992). Impairment of temporal organization of speech in basal ganglia diseases. *Brain and Language*, *43*, 386–399.

Whitfield, J. A., & Goberman, A. M. (2014). Articulatory-acoustic vowel space: Application to clear speech in individuals with Parkinson’s disease. *Journal of Communication Disorders*, *51*, 19–28.

Yule, G. (2006). *The Study of Language* (3rd ed.). Cambridge: Cambridge University Press.

Table 3.1. Fundamental frequency (f0), formant frequencies (F1-4), formant dispersion (D_f), and estimated vocal tract length (est. VTL) measured from the exhaled grunts of adult chacma baboons and geladas. Values listed as mean [SEM].

Species	Sex	Subjects (# per subject)	f0 (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	D_f (Hz)	est. VTL (cm)
chacma	male	9 (2-15)	55.442 [3.033]	285.041 [7.716]	1041.743 [34.095]	1855.466 [37.803]	2786.111 [67.944]	833.69 [22.795]	21.121 [0.594]
chacma	female	15 (1-5)	130.61 [5.553]	400.38 [12.827]	1339.255 [22.534]	2528.91 [77.114]	3315.737 [69.447]	971.786 [23.467]	18.148 [0.416]
gelada	male	17 (1-23)	268.605 [7.945]	633.419 [12.212]	1512.539 [33.355]	2644.149 [60.05]	3739.164 [78.711]	1035.248 [27.013]	17.118 [0.519]
gelada	female	32 (1-6)	383.772 [9.946]	773.731 [24.192]	2037.829 [76.625]	3646.951 [94.151]	5220.054 [101.858]	1482.108 [30.07]	11.965 [0.255]

Table 3.2. Fundamental frequency (f0) and formant frequencies (F1-4) measured from the derived calls of adult male geladas. Values listed as mean [SEM].

Call type	Subjects (# per subject)	f0 (Hz)	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)
inhaled grunt	14 (1-19)	259.694 [16.64]	607.69 [16.746]	1473.629 [28.858]	2701.323 [62.49]	3766.168 [78.262]
exhaled moan	8 (1-6)	294.225 [14.119]	624.123 [34.747]	1644.51 [50.437]	2676.738 [85.385]	3778.811 [115.865]
inhaled moan	15 (1-8)	202.027 [21.293]	570.885 [19.799]	1494.908 [38.862]	2707.291 [66.193]	3873.213 [99.126]
wobble	9 (1-5)	226.81 [25.475]	593.682 [24.32]	1305.587 [46.38]	2261.6 [140.619]	3246.065 [176.438]
yawn	17 (1-2)	369.45 [33.448]	693.313 [25.637]	1333.182 [23.775]	2320.329 [67.317]	3156.244 [101.521]

Table 3.3. Results of five LMMs used to test for differences in mean fundamental (f0) and formant (F1-4) frequencies (Hz) between exhaled grunts and all derived call types. Each column represents a single LMM.

Derived call type	f0			F1			F2			F3			F4		
	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**
Inhaled grunts	-5.12 [28.15]	0.18	0.8564	-25.68 [26.85]	0.96	0.3435	-38.91 [46.39]	0.84	0.4043	66.65 [95.89]	0.70	0.4899	50.09 [130.45]	0.38	0.7026
Exhaled moans	17.89 [34.00]	0.53	0.6006	-14.44 [32.38]	0.45	0.6573	131.97 [55.11]	2.40	**0.0192**	42.02 [115.15]	0.37	0.7164	32.67 [156.92]	0.208	0.8358
Inhaled moans	-69.30 [27.81]	2.49	**0.0155**	-63.54 [26.51]	2.40	**0.0201**	-17.63 [45.53]	0.39	0.6997	65.00 [94.49]	0.69	0.4942	144.52 [128.64]	1.12	0.2662
Wobbles	-44.75 [33.23]	1.35	0.1824	-39.94 [31.58]	1.27	0.2104	-206.95 [52.98]	3.91	**0.0002**	-383.58 [111.72]	3.43	**0.0010**	-488.94 [152.54]	3.21	**0.0021**
Yawns	99.33 [26.92]	3.69	**0.0005**	57.96 [25.66]	2.26	**0.0280**	-179.36 [44.08]	4.07	**0.0001**	-316.42 [91.46]	3.46	**0.0010**	-571.43 [124.51]	4.59	**< 0.0001**

* Bolded values indicate P < 0.05

Table 3.4. Results of five LMMs used to test for differences in the bandwidth of fundamental (f0) and formant (F1-4) frequencies (Hz) between exhaled grunts and all derived call types. Each column represents a single LMM.

Call type	f0			F1			F2			F3			F4		
	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**	Estimate [SEM]	*t*	*P**
Inhaled grunts	-11.61 [26.94]	0.43	0.6681	-0.81 [44.87]	0.02	0.9856	-75.427 [80.54]	0.94	0.3521	-40.37 [87.48]	0.46	0.6460	-41.73 [97.68]	0.43	0.6708
Exhaled moans	40.81 [32.41]	1.26	0.2126	70.55 [53.31]	1.32	0.1898	92.16 [95.68]	0.96	0.3386	55.66 [104.51]	0.53	0.5961	192.28 [117.12]	1.64	0.1055
Inhaled moans	1.25 [26.57]	0.05	0.9625	42.29 [44.05]	0.96	0.3401	-42.07 [79.06]	0.53	0.5963	7.13 [86.03]	0.08	0.9342	108.47 [96.20]	1.13	0.2639
Wobbles	3.77 [31.52]	0.12	0.9051	205.46 [51.26]	4.01	**0.0001**	255.99 [92.00]	2.78	**0.0068**	216.48 [100.89]	2.15	**0.0354**	112.05 [113.45]	0.99	0.3268
Yawns	308.87 [25.71]	12.01	< **0.0001**	243.90 [42.65]	5.72	**< 0.0001**	406.03 [76.55]	5.30	**< 0.0001**	425.91 [83.28]	5.11	**< 0.0001**	643.97 [93.11]	6.92	**< 0.0001**

* Bolded values indicate P < 0.05

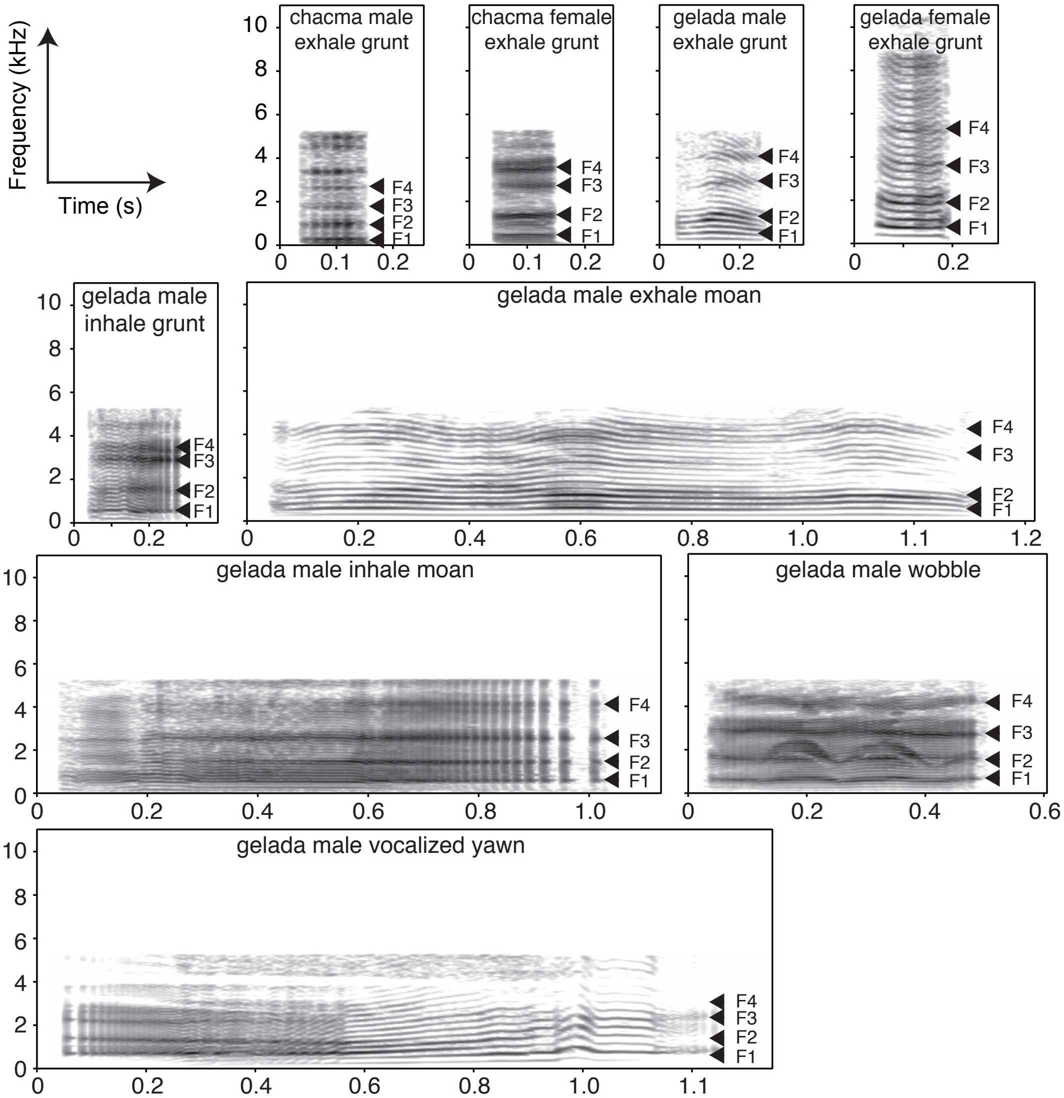


Figure 3.1. Spectrograms (Hanning window of 25 ms) of adult male and female chacma baboon and gelada exhale grunts, as well as a gelada male derived call types (inhaled grunt, exhaled moan, inhaled moan, wobble, and yawn). These are call types produced during affiliative contexts. Mean formant (1-4) frequencies are denoted by arrows. Spectrograms were made in Praat.

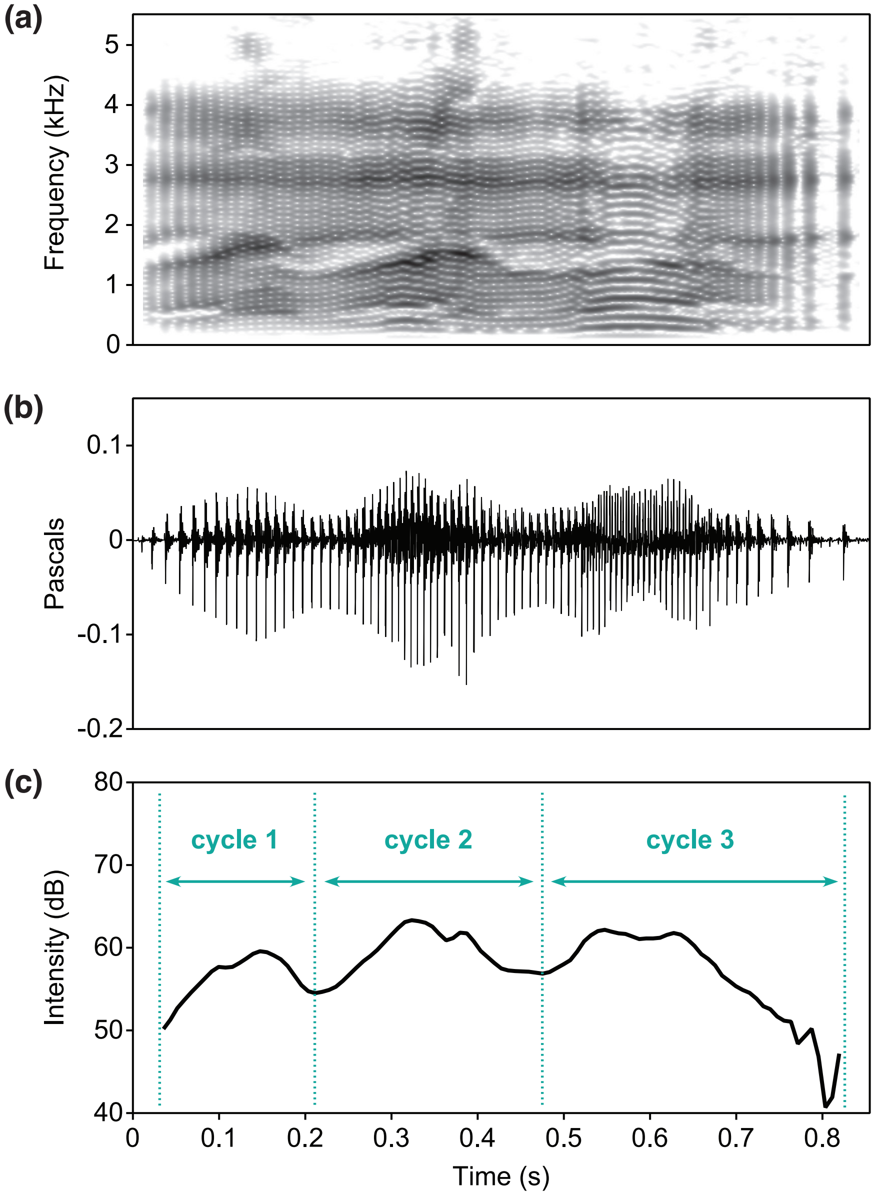


Figure 3.2. A gelada male wobble: (a) spectrogram (Hanning window of 25 ms), (b) waveform, and (c) intensity contours used to extract cycles start/end times (i.e., the local amplitude minimums of intensity tiers). Images were produced in Praat.

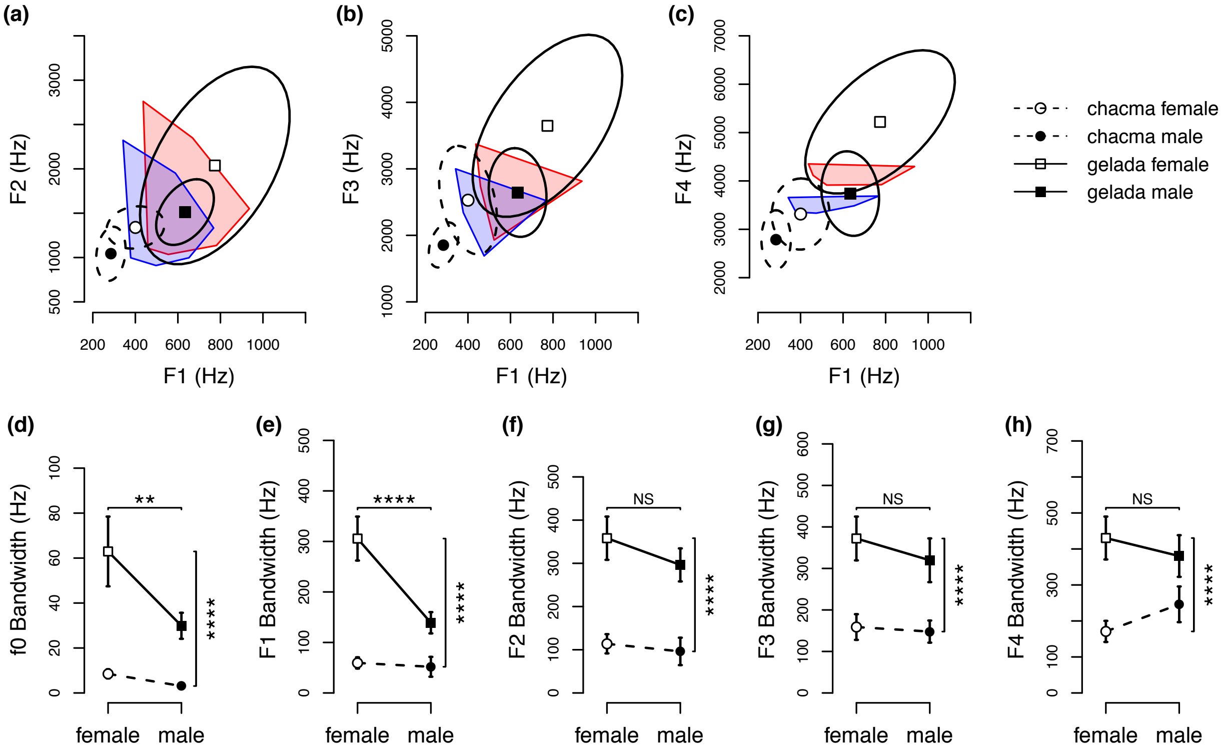


Figure 3.3. Between-species comparisons of (a-c) formant space, (d) fundamental frequency (f0) modulation and (e-h) formant modulation. Formant space comparisons were made for (a) F1-F2 space, (b) F1-F3 space, and (c) F1-F4 space in chacma baboon and gelada exhaled grunts and in human English vowels. Circle symbols and dotted lines (chacma baboons) and squares and solid lines (geladas) represent the mean formant frequencies surrounded by the 95% confidence interval ellipses. Open symbols represent females and solid symbols represent males. Human vowels (red – women; blue –men) are represented by convex hulls determined from published vowel formant frequencies (Hillenbrand, Getty, Clark, & Wheeler, 1995). f0 and formant bandwidths are used to characterize modulation (d-h) in chacma baboon and gelada exhaled grunts. Points and whiskers indicate mean ± 2 SEM. To make comparisons, ANOVAs included species, sex and the species*sex interaction as fixed effects. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. Sample sizes (each data point representing a single individual) are reported in Table 1.

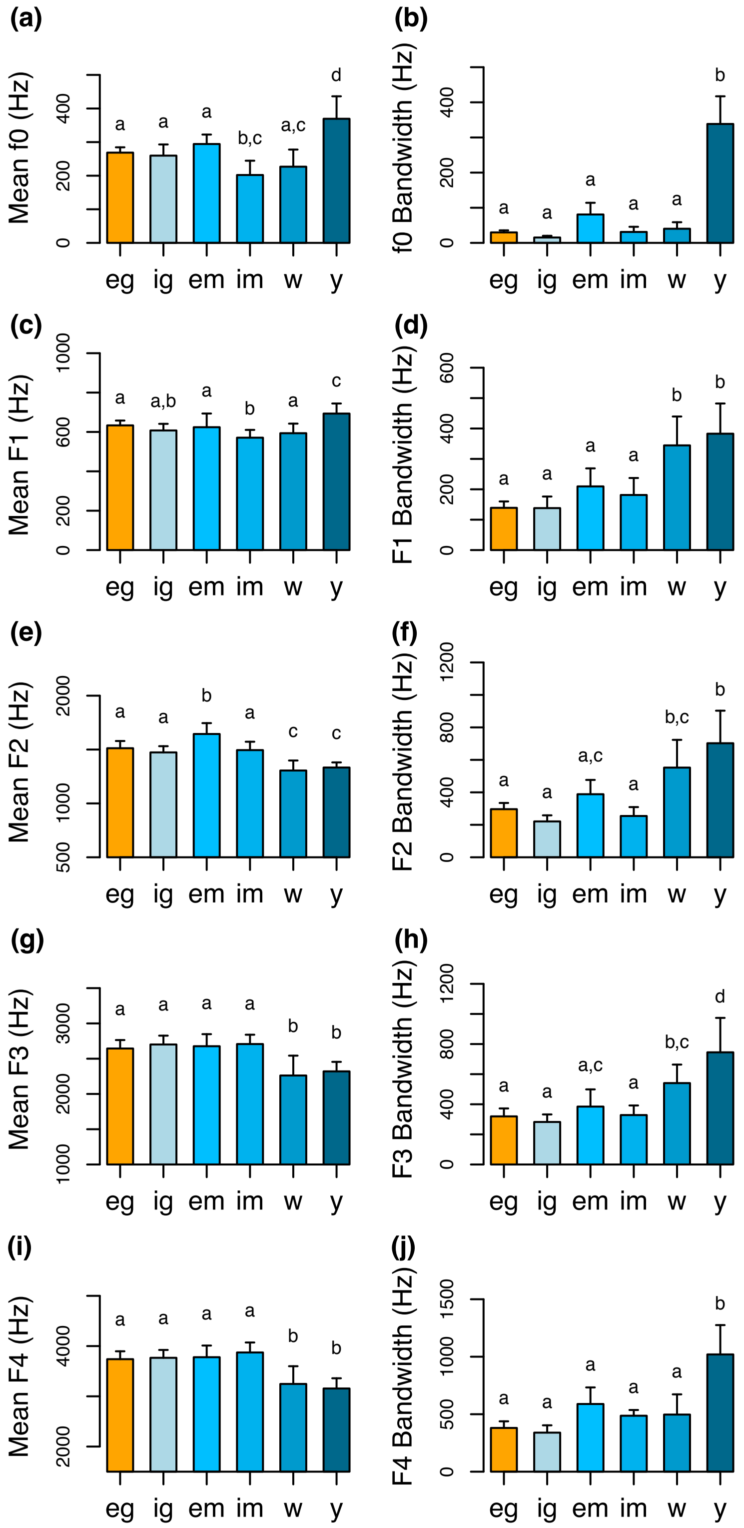


Figure 3.4. Comparisons of the mean frequencies and bandwidths of (a-b) f0, (c-d) F1, (e-f) F2, (g-h) F3, and (i-j) F4 frequencies of gelada male exhaled grunts (orange) and derived call types (shade of blue). The six call types include exhaled grunts (“eg”), inhaled grunts (“ig”), exhaled moans (“em”), inhaled moans (“im”), wobbles (“w”), and

yawns (“y”). Bars and whiskers indicate mean ± 2 SEM. To make comparisons, LMMs included f0 and F1-F4 means or bandwidths as the dependent variable, call type as the fixed effect, and caller identity as the random effect. Letters above the bars represent the outcome of pairwise comparisons between the call types. Calls that do not share a letter differed from each other at the level of $P < 0.05$. Sample sizes (each data point representing a single individual) are reported in Tables 1 and 2.

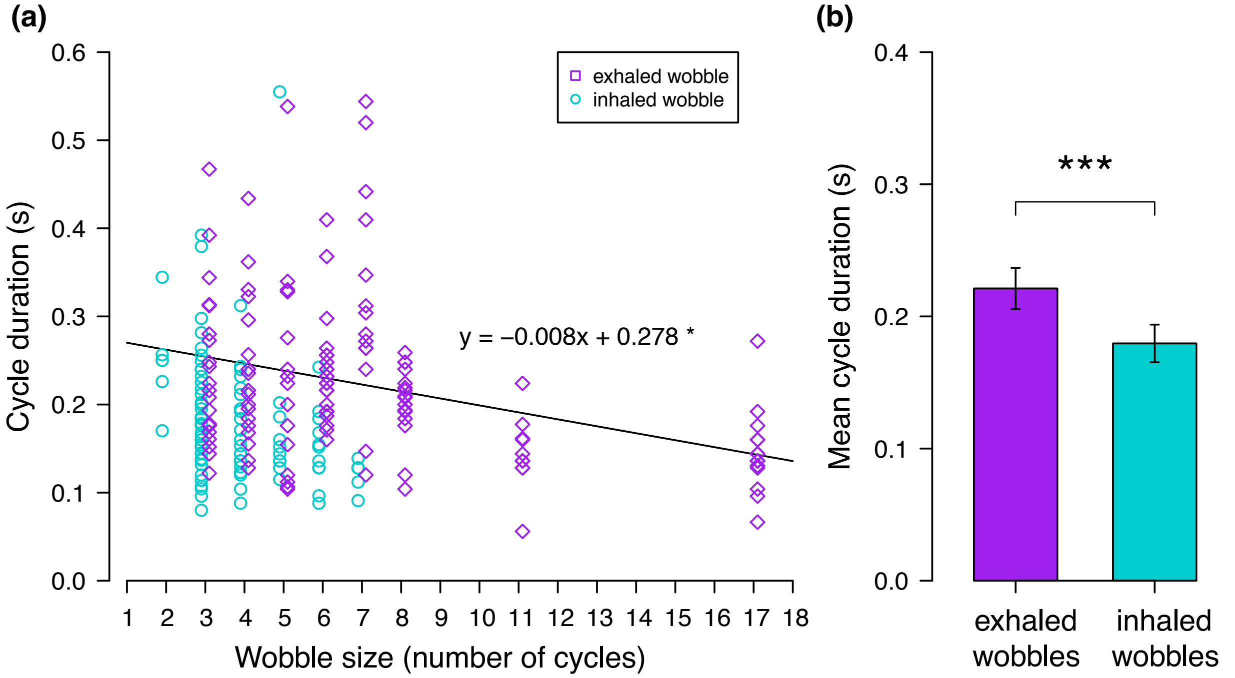


Figure 3.5. Associations between cycle duration and (a) wobble size and (b) production mode. Cyan circles represent individual cycles from inhaled wobbles and purple diamonds represent individual cycles from exhaled wobbles. Points and whiskers in plot (b) indicate mean ± 2 SEM. To make comparisons, LMMs included cycle duration as the dependent variable, wobble size and production mode as fixed effects, and wobble recording, cycle position, male identity as random effects. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$.

Chapter 4:

Gelada vocal sequences follow Menzerath's linguistic law

Abstract

Identifying universal principles underpinning diverse natural systems is a key goal of the life sciences. A powerful approach in addressing this goal has been to test whether patterns consistent with linguistic laws are found in non-human animals. Menzerath's law is a linguistic law that states that the larger the construct, the smaller the size of its constituents. Here, we present the first evidence that Menzerath's law holds in the vocal communication of a non-human species. We show that in vocal sequences of wild male geladas (*Theropithecus gelada*), construct size (sequence size in number of calls) is negatively correlated with constituent size (duration of calls). Call duration does not vary significantly with position in the sequence, but call sequence composition does change with sequence size and most call types are abbreviated in larger sequences. We also find that inter-call intervals follow the same relationship with sequence size as do calls. Finally, we provide formal mathematical support for the idea that Menzerath's law reflects compression – the principle of minimizing the expected length of a code. Our findings suggest that a common principle underpins human and gelada vocal communication, highlighting the value of exploring the applicability of linguistic laws in vocal systems outside the realm of language.

Introduction

Identifying fundamental principles that underpin diverse natural phenomena is a central goal of the life sciences (1, 2). The existence of such principles, revealed by the occurrence of common statistical patterns, can shed light on the basic processes shaping biological systems (3). In recent years, exploration of the universality of the statistical laws of human language has proved a fruitful starting point for identification and investigation of these fundamental principles (4, 5). The power of such an approach is illustrated by recent studies of the generality of Zipf's law of abbreviation (6). This linguistic law, also commonly known as Zipf's law of brevity, states that more frequently used words tend to be shorter, and it has been found to hold true in all languages assessed to date (5). Analyses of the non-vocal surface behavioral repertoire of dolphins (7), the vocal repertoire of Formosan macaques (8), close-range calls of common marmosets (9), and social calls of four species of bats (10) reveal that they too conform to an inverse general relationship between magnitude (*e.g.*, duration) and frequency of use. This common pattern provides evidence that compression – the information theoretic principle of minimizing the expected length of a code – is a general principle of animal (including human) behavior, reflecting selection for energetic efficiency (4, 11).

Further evidence of compression in human language - this time not at the level of individual elements, but at the level of elements combined into sequences - may come from studies of another linguistic law, Menzerath's law, which states simply that "the greater the whole, the smaller its constituents" (12–14). Traditionally, this law has been used to explore the structuring of language in written text, and evidence supports this law at different scales of analysis: the longer a word in terms of syllable number, the shorter on average are its constituent syllables (*e.g.*, 14), and the longer a sentence in terms of number of clauses, the shorter on average those clauses are (*e.g.*, 16). Although Menzerath's law was originally induced as a linguistic law, it has since been applied to a wide range of systems beyond human language: a negative correlation between construct and constituent size has been found in domains such as music (16) and at the molecular level, with species that have a larger number of chromosomes tending to have a smaller mean chromosome size (4, 17–20), genes that have a higher number of exons having shorter exons on average (21, 22), and proteins with a higher number of domains having on average shorter domains (23). This broad adherence to Menzerath's law can be interpreted as a manifestation of self-organization (24) and compression of information (14, 25), and suggests that these processes may be widespread in shaping multi-level systems as diverse as macromolecules, language and music.

To date, no study has tested Menzerath's law in the vocal communication of any species except our own. Carrying out such studies is essential if the generality of this law across different communication systems is to be evaluated. Such an investigation would also contribute to our understanding of the multifaceted nature of animal vocal signaling; identifying the processes by which animals take singular sounds and combine them into diverse sequences is a key goal for those wishing to quantify, compare and explain variation in vocal complexity across taxa, including humans (26). Moreover, finding conformity to Menzerath's law in animal vocal behavior would highlight an important commonality between human language and the vocal systems of other animals, with respect to the basic structural patterns underpinning how sounds are combined into larger units and structures. This commonality would provide evidence for selection acting on the communicative systems of these disparate taxa, either to constrain them to a common ancestral state, or to drive their convergent evolution (5).

Here we test Menzerath's law in the vocal communication of wild geladas (*Theropithecus gelada*), a gregarious primate species with a well described vocal repertoire, and in which individuals - typically adult males - produce long and complex vocal sequences during affiliative contexts (27–30). The constituent parts of these sequences are discrete, individual calls, and six different call types in total are seen in sequences. Playback experiments indicate that sequences composed of only the most common call type - grunts - elicit weaker responses from receivers than do more varied sequences; furthermore, regardless of composition, playbacks of sequences comprised of multiple call types elicit similar strong responses from receivers (31). These results suggest that, in contrast to what happens when words are combined to form sentences, geladas do not seem to combine elements into sequences to convey a different meaning. Instead, these vocal sequences may function similarly to bird song in that increasing the diversity of calls within them provides a more effective way to convey the same message (32).

First, we test for (a) a negative correlation between the number of calls in a sequence and the average duration of these calls; this provides a test of Menzerath's law in the gelada vocal system. Patterns consistent with Menzerath's law could result from one or more different processes, which we explore in further analyses. One such process would be energetic or breathing constraints on vocal production leading to shorter vocal utterances being made later in sequences (and more of these shortened calls being seen in sequences with more calls overall); we therefore assess whether (b) durations of calls become shorter later in sequences. Patterns consistent with the law could also occur if sequences with more calls contain a higher proportion of the shorter call types in the repertoire; therefore, we test whether (c) sequence composition changes with the number of calls. Another possible process underlying emergence of this law would be calls of a particular type being shortened in duration in sequences with more calls; we explore this possibility by testing whether (d) within specific call types, durations are shorter in sequences containing a larger number of calls. We then extend our analyses to explore the silences between calls (inter-call intervals) testing whether (e) Menzerath's law also holds for silences, such that inter-call intervals are shorter in sequences with more calls, and (f) durations of inter-call intervals become shorter later in sequences. Finally, as a link between Menzerath's law and the principle of compression has previously been proposed (14, 25), but not systematically established, we (g) use a mathematical approach to provide formal support for this link and (h) put forward a unified explanation for the origins of both Menzerath's law and Zipf's law of abbreviation.

Materials and Methods

(a) Study site and subjects

Data for this study come from 57 adult male geladas across three different bands in one wild community (about 1200 individuals) living in the Sankaber area of the Simien

Mountains National Park, Ethiopia (64). Gelada bands are comprised of smaller harem-like units with a leader male, 0-3 follower males, and 1-11 females with their immature offspring (64). Male geladas were chosen as the target of the study because they produce more complex vocalizations than females (27, 30). The study males included 33 unit leaders, 14 followers, and 7 males who were a leader and a follower at different times during the study period, from across 30 reproductive units. The remaining 3 study males were bachelors from all-male groups in the study bands, and one of these males was a leader and a bachelor at different times during the study period (65). This population has been under intensive behavioral study since January 2006, and study subjects are fully habituated to human observers on foot (approach distances less than 3 m) (66).

(b) Vocal sequence data collection and processing

From March 2008 to June 2014, we opportunistically recorded vocal sequences during behavioral observations, using a Sennheiser ME66 directional microphone connected to a Marantz PMD 660 or 661 Digital Recorder. We define a vocal sequence (*i.e.*, the whole construct) as one or more discrete calls (*i.e.*, the constituent parts) made by the same individual and, for those sequences with two or more calls, separated by an inter-call interval of less than 5 seconds (Figure 4.1). Sequence size refers to the number of calls in the sequence. Call (or interval) position refers to the placement of the call (or interval) in the entire sequence (1st, 2nd, 3rd, 4th, etc.). To control for behavioral context, we focus only on the vocal sequences made during close-range affiliative social situations. These ‘contact call’ vocal sequences are thought to play an important role in facilitating contact with group members, and they are known to vary widely in size, composition, and complexity of individual calls (27, 30). Six call types are made during these vocal sequences: exhale grunts, exhale moans, inhale grunts, inhale moans, wobbles and yawns (30). Call types in a vocal sequence were described at the time of the recording and reassessed during visual and auditory inspection of the spectrogram; an earlier study of gelada vocalizations, which examined calls from many of the sequences analyzed here, found 96% inter-observer reliability of call type identification (30). Because grunts and moans (an elongated version of a grunt) grade into each other, density plots of call durations (log-transformed) for exhaled and inhaled grunts/moans were used to determine thresholds to distinguish between grunt and moan call types. These thresholds were defined as the minimum value between the first two peaks. By this definition, the threshold between exhaled grunts and moans was 0.768 s and the threshold between inhaled grunts and moans was 0.513 s.

Our analyses focus on vocal sequences with sufficiently high noise-to-signal ratio to categorize call types, and which are uninterrupted by a call from another adult unit member. We used Avisoft (v. 5.1.12, R. Specht, Berlin) to generate spectrograms with a fast Fourier transformation (size of 1024 points) and to label the start time, end time and call type of all calls in a vocal sequence. These start and end times were used to calculate the call duration and inter-call interval duration. Calculations of call durations, inter-call interval durations, and all statistical tests were carried out in R 3.2.2. (67).

(c) Statistical analysis

To explore the relationship between call duration (and inter-call interval duration) and sequence size, we first used Spearman rank correlation tests, a method introduced to examine Menzerath's law and remain objective about the exact functional dependency between the variables (22, 37, 68). In these tests, the duration of individual calls or inter-call intervals for each sequence size was correlated with sequence size. While this can help to avoid potential problems of previous research on Menzerath's law (37), we recognize that its application with our dataset involves pseudoreplication. Therefore, linear mixed models (LMM) were constructed in order to assess the relationships between call (and interval) duration and sequence size while taking call (or interval) position, call type, male identity and sequence recording (*i.e.*, the vocal sequence that a call was from) into account. We ran the LMMs in R 3.2.2. using the function lmer of the R package lme4 (69). The lmerTest package was implemented to determine the significance of the LMM coefficients (70). (a) First, we examined variation in individual call duration by including sequence size, call position and call type as fixed factors and male identity and the sequence recording as random effects. (b) An effect of call position was assessed in the LMM for part (a), and to examine further the relationship between call position and individual call duration, we ran an additional LMM with duration of the first call in vocal sequences as the dependent variable, sequence size as a fixed effect, and male identity as a random effect. (c) To examine whether the proportion of call types changes as sequence size changes, we ran six LMM models, one for each call type. For these models, the proportion of each call type was calculated for all vocal sequences as the number of calls of a given call type divided by the sequence size. Call type proportion was included as a dependent variable, with sequence size as a fixed effect and male identity as a random effect. (d) Then, to assess whether a relationship between call duration and sequence size applied to specific call types, six LMM models were run, one for datasets limited to each of the call types. In these models, call duration was the dependent variable with sequence size as a fixed effect and with male identity and sequence recording as random effects. See Supporting Information for rationale behind the exclusion of call position as a fixed effect for the LMMs on datasets for specific call types and the results of analyses including call position. (e) We examined variation in inter-call interval duration by including sequence size, interval position and call type as fixed effects and male identity and sequence recording as random effects. (f) An effect of inter-call interval position was assessed in the LMM for part (e), and to examine further the relationship between interval position and interval duration, we ran an additional LMM with duration of the first inter-call interval as the dependent variable, sequence size as a fixed effect, and male identity as a random effect. All statistical tests were assessed using a significance level of α = 0.05.

Results

We analyzed 1065 vocal sequences (comprised of 4747 individual calls) recorded from 57 study males (1-113 sequences per male). The sequence sizes ranged between 1 and 26 calls (242 single call sequence recordings, 1-156 recordings for sequences with two or more calls, with 1-24 recordings per male per sequence size). There were no recordings available for sequences of 22 calls or 25 calls. Since 242 sequences were comprised of a single call, this means that 823 sequences were used in the analyses involving inter-call intervals. Total sequence duration averaged 1.97 s (SEM: 0.05 s; range: 0.12 s to 8.33 s), and the distribution of the sequences sizes was consistent with a geometric distribution, suggesting that the probability of sequence termination is not dependent on sequence size (see supporting information for further details about the variation of sequence duration and size). Example spectrograms of short and long vocal sequences from the same male, showing each of the six call types, are shown in Figure 4.1.

(a) Is call duration negatively correlated with sequence size?

To test whether there is a negative association between construct size and the duration of constituent parts, we tested for a correlation between sequence size and the duration of individual calls. Supporting Menzerath's law, individual call duration was negatively correlated with sequence size ($r_s = -0.229$, $n = 4747$, $p < 0.0001$ - Figure 4.2A). This finding was further supported by a Linear Mixed Model (LMM) that included individual call duration as the dependent variable, sequence size, call position, and call type as fixed effects, and male identity and sequence recording (*i.e.*, the vocal sequence that a call was from) as random effects ($P < 0.0001$ - Table 4.1).

(b) Does call duration get shorter later in the sequence?

To test whether such a Menzerath's law effect could be due to constraints on vocal production, resulting in shorter calls towards the end of a sequence, we explored the role of call position in the LMM (Table 4.1). Call position was not associated with call duration, indicating that calls were relatively similar in duration throughout an entire vocal sequence (Figure 4.2B). This finding was corroborated by an LMM showing that the duration of first position calls was negatively associated with sequence size (Estimate ± SEM: -0.020 ± 0.005, $t = 4.151$, $P < 0.0001$), indicating that the duration of calls in the beginning of sequences reflected sequence size.

(c) Do proportions of call types change with sequence size?

To test whether a Menzerath's law effect could be due to a varying proportion of shorter or longer call types in sequences of different sizes, we explored differences in duration across the six call types and examined associations between sequence size and proportions of each call type given in each sequence size. There was marked variation in duration among the six call types (Table 4.2). On average, inhale grunts were the shortest of the call types, exhale grunts were the second shortest call type, followed by inhale moans, wobbles, yawns and exhale moans (Table 4.1 and Table 4.2). The proportions of both exhale grunts (Estimate ± SEM: -0.019 ± 0.002, $t = 8.130$, $P < 0.0001$) and exhale moans (Estimate ± SEM: -0.008 ± 0.002, $t = 4.747$, $P < 0.0001$) in vocal sequences were negatively associated with sequence size (Figure 4.3). The proportions of inhale grunts (Estimate ± SEM: 0.023 ± 0.001, $t = 20.173$, $P < 0.0001$), inhale moans (Estimate ± SEM: 0.005 ± 0.001, $t = 3.979$, $P = 0.0001$) and wobbles (Estimate ± SEM: 0.001 ± 0.000, $t = 4.676$, $P < 0.0001$) in vocal sequences were positively associated with sequence size (Figure 4.3). The proportion of yawns in vocal sequences was unrelated to sequence size (Estimate ± SEM: -0.001 ± 0.001, $t = 1.351$, $P = 0.1772$ - Figure 4.3).

(d) Do calls of a particular type shorten in duration in longer sequences?

To test whether a Menzerath's law effect could be due to the shortening of specific call types in longer sequences, we explored the relationships between call duration and sequence size within particular call types using LMMs. Individual call duration was negatively associated with sequence size for exhale grunts (Estimate ± SEM: -0.002 ± 0.001, $t = 2.299$, $P = 0.0219$), exhale moans (Estimate ± SEM: -0.044 ± 0.013, $t = 3.471$, $P = 0.0006$), inhale grunts (Estimate ± SEM: -0.005 ± 0.001, $t = 6.001$, $P < 0.0001$) and wobbles (Estimate ± SEM: -0.083 ± 0.029, $t = 2.873$, $P = 0.0090$) (Figure 4.4). Call duration was unrelated to sequence size for inhale moans (Estimate ± SEM: 0.003 ± 0.003, $t = 0.863$, $P = 0.3892$) and yawns (Estimate ± SEM: 0.008 ± 0.010, $t = 0.771$, $P = 0.4422$) (Figure 4.4).

(e) Is inter-call interval duration negatively correlated with sequence size?

We extended our definition of constituent parts to include the inter-call intervals (the silences between calls) in vocal sequences and carried out analyses similar to those using individual call duration as a dependent variable. Supporting Menzerath's law, individual inter-call interval duration for each sequence was negatively correlated with sequence size ($r_s = -0.413$, $n = 3682$, $P < 0.0001$ - Figure 4.5A). This finding was further supported by a LMM that included individual inter-call interval duration as a dependent variable, sequence size and interval position as fixed effects, and male identity and sequence recording as random effects ($P < 0.0001$; Table 4.3).

(f) Do inter-call intervals become shorter later in a sequence?

To test whether such a Menzerath's law effect of inter-call intervals could be due to constraints on vocal production, resulting in shorter intervals towards the end of a sequence, we explored the role of interval position in the LMM (Table 4.3). There was support for the prediction that inter-call interval duration shortens as a sequence progresses, as interval duration was negatively associated with interval position (Figure 4.5B). Moreover, an LMM showed that the duration of first position inter-call intervals was negatively associated with sequence size (Estimate ± SEM: -0.021 ± 0.003, t = 6.365, $P < 0.0001$), indicating that the intervals in the beginning of sequences also reflected sequence size.

(g) A formal mathematical exploration of the link between Menzerath's law and compression

In quantitative linguistics, the presence of Menzerath's law has been interpreted as a compression effect ((26), p. 147; (15), p. 42); however, the connection between this law and the problem of compression in information theory is not obvious and has never been formally explored mathematically. In standard coding theory, the problem of compression is based on the minimization of the mean code length (33)

$$L = \sum_{i=1}^{V} p_i l_i, \qquad [1]$$

where p_i and l_i are, respectively, the probability and the length of the i-th element of a repertoire (*e.g.*, an alphabet) of size V. Solving the problem of compression consists of finding the lengths so as to minimize L, assuming that the probabilities of the elements are constant and that two elements cannot have the same code. The minimization of L and a negative correlation between probability and length are intimately related (11). For instance, it has been shown that the minimization of L leads to the law of abbreviation, *i.e.* a negative correlation between probability and length (4, 11). For Menzerath's law in human language and the vocalizations of geladas, we propose the minimization of the following function

$$M = \sum_{i=1}^{T} l_i, \qquad [2]$$

where l_i is now the magnitude of the *i*-th construct in a series of *T* constructs. For instance, l_i could be the total duration of the *i*-th vocalization of a gelada or the length in words of the *i*-th sentence in text (see (5) for a review of costs associated to duration). Notice that there is an important difference between *L* and *M*; to illustrate this point, we borrow the concept of type and token from quantitative linguistics (34). While *L* is defined as a summation over types (elements of the repertoire or vocabulary), *M* is defined as a summation over tokens, namely occurrences of elements of the repertoire. Interestingly, *M* can be expressed equivalently as the total sum of the duration of the parts of every occurrence, namely

$$M = \sum_{i=1}^{T} \sum_{j=1}^{n_i} l_{ij}, \qquad [3]$$

where l_{ij} is the length of the *j*-th part of the *i*-th occurrence (token). If one assumes that given a certain occurrence, all the parts are identically distributed, it turns out that the expectation of *M* is

$$E[M] = \sum_{i=1}^{T} n_i E\left[l_{ij} \middle| i\right], \qquad [4]$$

where $E[l_{ij}|i]$ is the expected magnitude of the parts of the *i*-th occurrence, assuming that parts are identically distributed for a given occurrence. Now it is easy to see that *L* and E[*M*] follow the same scheme: both are a weighted sum of the magnitude of elements from a set. Applying the arguments employed to derive the law of abbreviation from the minimization of *L* (11) one can also derive Menzerath's law from the minimization of

E[*M*]. This argument relies on the simplifying assumption that parts are identically distributed, which is supported by the absence of an overall effect of call position (Table 4.1).

(h) A unified explanation for the origins of Zipf's law of abbreviation and Menzerath's law through a general cost function.

The generality of the principle of compression becomes more evident when regarding the minimization of *L* and E[*M*] as instances of the minimization of a generalized cost function

$$K = \sum_i x_i y_i, \qquad [5]$$

where x_i and y_i are respectively the weight and the magnitude of the *i*-th element of a set (we assume that all "x_i"s and "y_i"s are positive). *L* is a particular case of *K* when x_i and y_i are, respectively, the probability and the length of the *i*-th element (in this case, *K* is defined over the alphabet or species repertoire). E[*M*] is a particular case of *K* when x_i is the number of parts (the number of calls) of the *i*-th occurrence (token) and y_i is the expected magnitude of its parts; *i.e.*, $E[l_{ij}|i]$ (in that case, the set consists of occurrences or tokens). Applying the same arguments of Ferrer-i-Cancho et al. (11) to the minimization of *K*, a negative correlation between the weight of a unit and its magnitude is expected. To conclude, the minimization of *K* can be viewed as a general principle of compression, shedding light on the origins of both the Zipf's law of abbreviation and Menzerath's law. This provides theoretical support for the intuition that Menzerath's law is an effect of compression (14, 23, 25). Indeed, one can also conclude that minimization of *K* and a negative correlation between the weight of a unit and its magnitude (Zipf’s law of abbreviation, Menzerath’s law) are intimately related, based upon previous arguments for *L* (11).

Discussion

The complex vocal sequences of adult male geladas follow Menzerath’s law, with sequences that are larger - *i.e.*, that contain more calls - being composed of calls with shorter duration. This finding provides new support for the generality of this linguistic law beyond the realm of human language. There is currently great controversy about why diverse biological systems exhibit Menzerath’s law and related laws (4, 5, 8, 18, 20, 35), and indeed about whether such laws hold at all (9, 36). Our findings make a significant contribution to this ongoing debate, providing the first evidence for Menzerath’s law in non-human animal vocal systems, and complementing a growing body of comparable evidence from a range of mammalian species (7, 8, 10) that behavior can be described by another, related linguistic law – Zipf’s law of abbreviation (6). A virtue of our analysis of the presence of Menzerath’s law is that it considers the duration of individual calls within a sequence and thus the finding of the law cannot be attributed to an artefact of employing mean durations (37). Our results also provide new support for the hypothesis that Menzerath’s law, as is the case for Zipf’s law of abbreviation, is an effect of compression – the information theoretic principle of minimizing the expected length of a code.

The finding that gelada vocal sequences show the same negative relationship between the size of the whole construct and the size of its constituents, as is found in human language, suggests that equivalent principles of self-organization (24) underpin the vocal communication of our own species and another primate. While there are elementary differences between the vocal faculties of humans and of other animals (38–40), exploring and comparing mathematical, structural properties of their communication systems can be informative (26, 41). As language is inherently sequence-based and animals of many taxa, from bacteria (42) to great apes (43), combine individual signals into sequences, identifying basic patterns of sequence structure that are shared by human and non-human animal communication provides evidence for evolutionary preservation, or convergent evolution, of the processes underlying the emergence of such patterns (41). Importantly, adherence to Menzerath's law need not involve any cognitively demanding planning by the animals producing sequences of sound. Finding Menzerath's law in gelada communication suggests that strings of sound following this law could pre-date the evolution of meaningful combinations.

Our further analyses revealed insights into *how* the negative relationship between call duration and sequence size arises in gelada vocal sequences. Firstly, our analyses of call duration and call position in a vocal sequence suggest that conformity to Menzerath's law cannot be explained simply by physiological or mechanical constraints on vocal production leading to shorter calls being made at the end of sequences (and more of these abbreviated calls being added to the end of longer sequences). Considering all call types together, a negative association was not seen between call position and duration as would be expected if such constraints were important; this indicates that from the start of vocal sequences, male geladas make calls of the 'appropriate' duration for that sequence size, and call duration does not then reduce predictably as the sequence progresses.

The emergence of patterns consistent with Menzerath's law may be in part due to the composition of vocal sequences varying with the number of calls. In particular, the proportion of exhale grunts decreased and the proportion of inhale grunts increased in larger sequences. This is likely to be due to males increasing the frequency with which they make both an inhale and an exhale grunt on the same breath as sequence size increases. Most significantly, once sequences exceed about 15 calls, inhale and exhale grunts - the two call types with the shortest duration - made up the vast majority of the constituent calls, and the relative proportion of each call type in the sequences varied little. Abbreviation of some individual call types in larger sequences may also underpin conformity to Menzerath's law: the durations of inhale grunts and exhale moans (the second and third most common call types in sequences) were negatively related to the number of calls in the sequence. There is little evidence from studies of human language of a comparable effect (*i.e.*, that the *same* constituent of a construct decreases in length in larger constructs), not least because Menzerath's law is almost exclusively investigated in written texts where a given word has a fixed length. However, one of the earliest studies of this phenomenon (conducted before its formal description as a linguistic law) suggests a similar shortening of a particular spoken sound when it appears in a larger construct. Grégoire (44) found that in spoken French, the duration of the pronunciation of the syllable *pâ-* decreased with increasing duration of the word or phrase that it began; the sound lasted 325ms in *pâte*, but only 195ms in *pâteuse*.

Conformity to Menzerath's law in gelada vocal sequences may reflect constraints linked to the respiratory and energetic demands of signal production. The lack of a strong association between call duration and position in the sequence does not preclude the possibility that physiological or mechanical constraints are important limiting factors in gelada vocal sequences. Non-human primates' relatively limited breathing control while vocalizing restricts their ability to produce long, continuous vocal sequences (45, 46), and an increase in the frequency with which vocalizations are made on both the inhale and exhale of the same breath could lead to hyperventilation (47). Additionally, as revealed both by comparative analyses (48) and single species studies (49–51), the duration of vocal signals in vertebrates can be constrained by energy availability. While reducing signal duration can therefore save energy and/or may be a necessary result of breathing patterns, it remains an open question as to whether or not the shortening of a vocalization results in the loss of transmission fidelity. Theoretical analyses of communication indicate that transmission fidelity is increased by adding redundancy, *e.g.* increasing signal duration (52), and studies of human speech support this idea: when asked to speak particularly clearly, people significantly increase the duration of individual speech sounds (53). A reduction in signal length also increases the risk both of confusing different signals and of perturbation of the signal by noise in the environment (54). Menzerath's law may therefore emerge in gelada vocal sequences as a compromise between effectiveness of communication, and the energetic demands and breathing constraints of vocal production.

Our analyses extended the standard approach of exploring Menzerath's law by considering the gaps between constituents of a construct (inter-call intervals in our study), as well as the constituents themselves (calls); to our knowledge, no study of Menzerath's law in humans has explored these inter-constituent gaps. The same negative relationship with sequence size was found for inter-call interval duration as was seen for call duration. Thus, larger vocal sequences have a faster tempo – they contain calls of shorter duration, with shorter gaps in between, and therefore are delivered at a higher rate. This variation in sequence tempo may reflect the emotional arousal state of the vocalizing animal (55) and potentially signal such information to receivers (56). A promising avenue for future research will be to investigate whether these longer and faster sequences have a different communicative function from shorter and slower sequences, such as expressing affective state to potential listeners (55). Indeed, it has been proposed that the rhythm of gelada vocal sequences has important communicative function in itself (27), but this idea has never been formally tested. One advantage of longer sequences being produced at a faster tempo is that it reduces the risk of being 'talked over' by other geladas; this is a serious potential problem, due to a noisy environment of conspecific vocalizations in which these animals need to communicate (30). Notably, the duration of inter-call intervals was negatively related to position in the sequence; this was not the case for calls, where no such relationship was seen. This indicates that tempo increases towards the end of sequences.

Although no previous study of non-human animal communication has explicitly tested Menzerath's law, results from the literature indicate that a negative relationship between sequence size and call duration is not inevitable. For example, baboons give longer sequences of grunts in infant handling compared to movement contexts, but the grunts in the longer sequences were longer, not shorter, in duration (57). Similarly, in a study of Barbary macaque female copulation calling (58), sequences given around the time of ovulation had a higher number of calls than sequences given early in the cycle, and these calls were again longer in duration. The difference between these findings and the present results may be due to the nature of the respective vocal sequences. Unlike gelada vocal sequences, sequences of baboon grunts are both homogenous and shorter, and so there may be less opportunity for Menzerath's law to emerge. In Barbary macaques, copulation sequences not only consist of the same call type repeated multiple times, but are very loud and appear to be aimed at distant as well as nearby receivers (59). In long-range communication, a potential reason for Menzerath's law not to emerge is the conflict between compression and transmission success, which lead to opposing constraints; it is easier for compression to dominate when successful communication over long distances is not necessary (5). Our analysis of short-range vocalizations in geladas provides support for the hypothesis that compression prevails in short range communication (10).

Although a connection between Menzerath's law and compression has been suggested various times (14, 25), it has not been formally investigated. Here, we provide the first mathematical demonstration of the connection between compression and Menzerath's law. We also provide the first unified explanation for both Zipf's law of abbreviation and Menzerath's law by means of a general cost function, which could explain the recurrence of these patterns across biological systems and levels of organization. The power of the argument relies on its mathematical simplicity and its general scope: it covers both individual elements and also the (possibly recursive) sequential structures they form. Furthermore, our explanation makes no linguistic or communicative assumptions about the elements or their combinations (*e.g.*, with respect to whether they convey meaning, or serve a particular function). Therefore, our theoretical framework can bridge gaps across a broad range of biological disciplines.

Conclusion

Quantitative methods have proved extremely powerful in the field of linguistics, for example revealing shared fundamental properties of geographically and morphologically diverse languages, which are not revealed by qualitative approaches (60, 61). The full power of linguistic laws may, however, extend far beyond the field in which they were developed (4, 5). Testing for conformity to linguistic laws in other biological systems facilitates a move from exploration of universal properties of language (62) to the development and testing of hypotheses about the fundamental principles that may explain the repeated occurrence of these statistical patterns across diverse biological contexts (4, 5). Our demonstration of a pattern consistent with Menzerath's law in gelada vocal communication further highlights the value of exploring the full scope of linguistic laws outside the realm of language. This law can be tested wherever a structure, process or system can be broken down into a construct and its constituents; the law's mathematical simplicity belies its enormous potential explanatory power, and we hope our work will encourage others to test the generality of Menzerath's law in biology. Finally, our demonstration of how a generalized principle of compression produces Menzerath's law, in conjunction with our previous work indicating that Zipf's law of abbreviation also reflects this principle (5), suggests that compression may underpin biological information systems in a broad sense. Compression exists not only at different levels of the chemistry of life - codons (63), proteins (23), genes (22) and entire genomes (19) - but also in multiple forms of animal behavior, from elementary patterns of behavior (7) to vocal communication in non-humans (8, 10) and human language (4, 25).

Acknowledgements

This research and the gelada project was supported by grants from the National Geographic Society (Gr# 9122-12, W304-14 to MLG; Gr# 67250 to TJB), the Wildlife Conservation Society (SSF Gr# 67250 to TJB), the Leakey Foundation, the National Science Foundation (Gr# BCS-0715179, BCS-0962118 to TJB), University of Michigan, and Petridish (primary supporter: John Allen). RFC was funded by the grants 2014SGR890 (MACDA) from AGAUR (Generalitat de Catalunya) and also the APCOM project (TIN2014-57226-P) from MINECO (Ministerio de Economia y Competitividad). We are grateful to the Ethiopian Wildlife Conservation Authority and the wardens and staff of the Simien Mountains National Park for granting us permission to conduct this research in Ethiopia. The gelada research was approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan and adhered to the laws and guidelines of Ethiopia. We would like to thank Aliza le Roux, Marcela Benitez, Megan Gomery, Elizabeth Johnson and John Allen for assistance with collecting acoustic data in the field and Niki Abdou for helping us with acoustic analysis. We also thank Kerby Shedden and Josh Errickson at the University of Michigan Center for Statistical Consultation and Research (CSCAR) for their statistical advice. We thank members of the University of Roehampton Centre for Research in Evolutionary, Social and Interdisciplinary Anthropology (CRESIDA) for helpful comments on previous versions of this manuscript.

Supporting Information

1. Variation of vocal sequence duration and sequence size

To test whether there is an association between sequence size and sequence duration (the amount of time between the start of the first call and the end of the final call in a sequence), we carried out a Linear Mixed Model (LMM) that included sequence duration as a dependent variable, sequence size as a fixed effect, and male identity as a random effect. This analysis revealed that sequence size was positively associated with sequence duration (Estimate ± SEM: 0.326 ± 0.008, $t = 39.988$, $P < 0.0001$ - Figure 4.S1).

The empirical distribution of sequence sizes is close to a straight line when taking logs on the y-axis (Figure 4.S2), suggesting that sequence sizes follow a geometric distribution - *i.e.* the probability that a sequence has size l is $p(l) = \pi (1 - \pi)^{l-1}$ - where π is the probability of termination, the only parameter of the distribution. The maximum likelihood estimator of π, *i.e.* $1/L$, where L is the mean size of the sequences (71), gives $\pi = 0.224$. With this parameter, a geometric distribution provides a high quality visual fit to the actual distribution of sequence sizes (Figure 4.S2). The consistency of sequence sizes with a geometric distribution suggests that the process of generating sequences is 'memory-less' (72): during the process of sequence generation, the current length does not change the probability of ending the sequence.

2. Additional analyses of the relationships between the duration of particular call types and sequence size.

In the main text, call position is not included as a fixed effect in the analyses on call duration and sequence size for particular call types because inhale moans, exhale moans, wobbles and yawns are rarely produced repeatedly in a single vocal sequence. This means that our power of using repeated measures (*i.e.*, sequence recording as a random effect) is greatly reduced, limiting the ability of the Linear Mixed Models (LMM) to effectively compare calls at different positions within sequences of a given size. Also, this means that call position and sequence size are not always independent. Certain call types like yawns, inhaled moans and wobbles are frequently found near the end of vocal sequences. In these instances, 'sequence size' and 'call position' can be similar or the same for several of the data points. Consequently, results from the following LMMs (below) - in which call position is included as a fixed effect, and that we report for completeness sake - should be interpreted with caution.

As a supplement to our main analyses, we ran six LMMs (one for each call type) with call duration as a dependent variable, sequence size and call position as fixed effects, and male identity and sequence recording as random effects. When including call position as a fixed effect, the LMMs show that call duration is negatively associated with sequence size only in inhaled grunts and inhaled moans (Table 4.S1). The LMMs also show a negative association between call duration and call position for exhale grunts and exhale moans, and a positive association between call duration and call position for inhale grunts and inhale moans.

References

1. Poincaré H (1913) *The foundations of science: science and hypothesis, the value of science, science and method* (The Science Press, New York).
2. Popper K (1934) The logic of scientific discovery (Verlag von Julius Springer, Vienna).
3. von Bertalanffy L (1950) The theory of open systems in physics and biology. *Science (80-)* 111:23–29.
4. Ferrer-i-Cancho R, Forns N, Hernàndez-Feràndez A, Bel-enguix G, Baixeries J (2013) The challenges of statistical patterns of language: the case of Menzerath's law in genomes. *Complexity* 18:11–17.
5. Ferrer-i-Cancho R, et al. (2013) Compression as a universal principle of animal behavior. *Cogn Sci* 37:1565–1578.
6. Zipf G (1936) The psycho-biology of language: an introduction to dynamic philology (George Routledge and Sons Ltd., London).
7. Ferrer-i-Cancho R, Lusseau D (2009) Efficient coding in dolphin surface behavioral patterns. *Complexity* 14(5):23–25.
8. Semple S, Hsu MJ, Agoramoorthy G (2010) Efficiency of coding in macaque vocal communication. *Biol Lett* 6:469–71.
9. Ferrer-i-Cancho R, Hernàndez-Feràndez A (2013) The failure of the law of brevity in two new world primates. Statistical caveats. *Glottotheory* 4:45–55.
10. Luo B, et al. (2013) Brevity is prevalent in bat short-range communication. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 199:325–33.
11. Ferrer-i-Cancho R, Bentz C, Seguin C (2015) Compression and the origins of Zipf 's law of abbreviation. http://arxiv.org/abs/1504.04884.
12. Menzerath P (1954) *Die Architektonik des deutschen Wortschatzes* (Bonn, Dümmler).
13. Altmann G (1980) Prolegomena to Menzerath's law. *Glottometrika* 2:1–10.

14. Köhler R (2012) Quantitative syntax analysis (De Gruyter Mouton, Berlin).

15. Teupenhayn R, Altmann G (1984) Clause length and Menzerath's law. *Glottometrika* 6:127–138.

16. Boroda MG, Altmann G (1991) Menzerath's law in musical texts. *Musikometrika 3*, ed Boroda MG (Brockmeyer, Bochum), pp 1–13.

17. Wilde J, Schwibbe MH (1989) Organizationsformen von Erbinformation Im Hinblick auf die Menzerathsche Regel. *Das Menzerathsche Gesetz in Informationsverarbeitenden Systemen*, eds Altmann G, Schwibbe MH, Kaumanns W (G. Olms, Hildesheim), pp 92–107.

18. Ferrer-i-Cancho R, Forns N (2010) The self-organization of genomes. *Complexity* 15:34–36.

19. Hernández-fernández A, Baixeries J, Forns N, Ferrer-i-Cancho R (2011) Size of the whole versus number of parts in genomes. *Entropy* 13:1465–1480.

20. Baixeries J, Hernández-Fernández A, Ferrer-i-Cancho R (2012) Random models of Menzerath-Altmann law in genomes. *BioSystems* 107:167–73.

21. Li W (2012) Menzerath's law at the gene-exon level in the human genome. *Complexity* 17:49–53.

22. Nikolaou C (2014) Menzerath-Altmann law in mammalian exons reflects the dynamics of gene structure evolution. *Comput Biol Chem* 53:134–143.

23. Shahzad K, Mittenthal JE, Caetano-Anollés G (2015) The organization of domains in proteins obeys Menzerath-Altmann's law of language. *BMC Syst Biol* 9:44.

24. Köhler R (1987) System theoretical linguistics. *Theor Linguist* 14:241–258.

25. Cramer I (2005) The parameters of the Altmann-Menzerath law. *J Quant Linguist* 12:41–52.

26. Kershenbaum A, et al. (2016) Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biol Rev* 91:13–52.

27. Richman B (1987) Rhythm and melody in gelada vocal exchanges. *Primates*

28:199–223.

28. Richman B (1976) Some vocal distinctive features used by gelada monkeys. *J Acoust Soc Am* 60:718–24.

29. Aich H, Moos-Heilen R, Zimmermann E (1990) Vocalizations of adult gelada baboons (*Theropithecus gelada*): Acoustic structure and behavioural context. *Folia Primatol* 55:109–132.

30. Gustison ML, le Roux A, Bergman TJ (2012) Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philos Trans R Soc Lond B Biol Sci* 367:1847–59.

31. Gustison ML, Bergman TJ (2016) Vocal complexity influences female responses to gelada male calls. *Sci Rep* 6:19680.

32. Nowicki S, Searcy WA (2004) Song function and the evolution of female preferences: why birds sing, why brains matter. *Ann N Y Acad Sci* 1016:704–23.

33. Cover TM, Thomas JA (2006) *Elements of information theory* (Wiley, Hoboken, NJ). 2nd Ed.

34. Covington MA, McFall JD (2010) Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). *J Quant Linguist* 17:94–100.

35. Solé R V (2010) Genome size, self-organization and DNA’s dark matter. *Complexity* 16:20–23.

36. Bezerra BM, Souto AS, Radford AN, Jones G (2011) Brevity is not always a virtue in primate communication. *Biol Lett* 7:23–5.

37. Ferrer-i-Cancho R, et al. (2014) When is Menzerath-Altmann law mathematically trivial? A new approach. *Stat Appl Genet Mol Biol* 13:633–644.

38. Pinker S (1995) *The language instinct: the new science of language and mind* (Penguin, London).

39. Deacon T (1997) *The symbolic species: the co-evolution of language and the brain* (Norton, New York).

40. Fitch TT (2010) *The evolution of language* (Cambridge University Press, Cambridge).

41. Kershenbaum A, et al. (2014) Animal vocal sequences : not the Markov chains we thought they were. *Proc R Soc B Biol Sci* 281:20141370.

42. Scott-Phillips TC, Gurney J, Ivens A, Diggle SP, Popat R (2014) Combinatorial communication in bacteria: Implications for the origins of linguistic generativity. *PLoS One* 9:2–6.

43. Clay Z, Zuberbühler K (2011) Bonobos extract meaning from call sequences. *PLoS One* 6:e18786.

44. Grégoire A (1899) Variation de la durée de la syllabe en français. *Le Parol* 1:161–176.

45. MacLarnon AM, Hewitt GP (1999) The evolution of human speech: the role of enhanced breathing control. *Am J Phys Anthropol* 109:341–63.

46. MacLarnon AM, Hewitt GP (2004) Increased breathing control: Another factor in the evolution of human language. *Evol Anthropol* 13:181–197.

47. Hewitt G, MacLarnon A, Jones KE (2002) The functions of laryngeal air sacs in primates: A new hypothesis. *Folia Primatol* 73:70–94.

48. Gillooly JF, Ophir AG (2010) The energetic basis of acoustic communication. *Proc R Soc B Biol Sci* 277:1325–31.

49. Thomas RJ (2002) The costs of singing in nightingales. *Anim Behav* 63:959–966.

50. Ward S, Speakman JR, Slater PJB (2003) The energy cost of song in the canary, *Serinus canaria*. *Anim Behav* 66:893–902.

51. Noren DP, Holt MM, Dunkin RC, Williams TM (2013) The metabolic cost of communicative sound production in bottlenose dolphins (Tursiops truncatus). *J Exp Biol* 216:1624–1629.

52. Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423.

53. Picheny MA (1986) Speaking clearly for the hard of hearing II Acoustic characteristics of clear and conversational speech. *J speech, Lang Hear Res* 29:434.

54. Brumm H, Zollinger SA (2011) The evolution of the Lombard effect: 100 years of psychoacoustic research. *Behaviour* 148:1173–1198.

55. Marler P (1977) The signaling behavior of apes. *How Animals Communicate*, ed Sebeok T (Indiana University Press, Bloomington, IN), pp 955–1035.

56. Seyfarth RM, et al. (2010) The central importance of information in studies of animal communication. *Anim Behav* 80:3–8.

57. Rendall D (2003) Acoustic correlates of caller identity and affect intensity in the vowel-like grunt vocalizations of baboons. *J Acoust Soc Am* 113:3390–3402.

58. Semple S, McComb KE (2000) Perception of female reproductive state from vocal cues in a mammal species. *Proc Biol Sci* 267:707–712.

59. Semple S (1998) The function of Barbary macaque copulation calls. *Proc R Soc B Biol Sci* 265:287–291.

60. Testitelova (1992) *Quanitative linguistics* (John Benjamins Publishing Company, Amsterdam).

61. Rasinger SM (2013) Quantitative research in linguistics: an introduction (Bloomsbury, London). 2nd Ed.

62. Evans N, Levinson SC (2009) The myth of language universals: language diversity and its importance for cognitive science. *Behav Brain Sci* 32:429–92.

63. Naranan S, Balasubrahmanyan VK (2000) Information theory and algorithmic complexity: Applications to linguistic discourses and DNA sequences as complex systems. *J Quant Linguist* 7:129–151.

64. Snyder-Mackler N, Beehner JC, Bergman TJ (2012) Defining higher levels in the multilevel societies of geladas (*Theropithecus gelada*). *Int J Primatol* 33:1054–1068.

65. Pappano DJ, Snyder-Mackler N, Bergman TJ, Beehner JC (2012) Social "predators" within a multilevel primate society. *Anim Behav* 84:653–658.

66. Bergman TJ (2010) Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proc R Soc B Biol Sci* 277:3045–53.

67. R Development Core Team (2015) R: A language and environment for statistical computing. Available at: http://www.r-project.org/.

68. Baixeries J, Hernández-Fernández A, Forns N, Ferrer-i-Cancho R (2013) The Parameters of the Menzerath-Altmann Law in Genomes. *J Quant Linguist* 20:94–104.

69. Bates D, Maechler M, Bolker B (2012) lme4: Linear mixed-effects models using S4 classes. Available at: http://cran.r-project.org/ package=lme4.

70. Kuznetsova A, Brockhoff PB, Christensen RHB (2013) lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Available at: http://cran.r-project.org/package=lmerTest.

References – Supporting Information

71. Ramachandran KM, Tsokos CP. The method of maximum likelihood. Mathematical statistics with applications. Boston, MA: Elsevier Inc.; 2009. p. 235–45.

72. DeGroot MD. The negative binomial distribution. Probability and Statistics. Reading, MA: Addison-Wesley; 1986. p. 258–63.

Table 4.1. Results of LMM used to test for relationships between call duration and sequence size, call position and call type.

Factor	Estimate	SEM	*t*	P[b]
Intercept	0.333	0.009	35.228	**< 0.0001**
Sequence size	-0.006	0.001	-5.041	**< 0.0001**
Call position	0.003	0.001	1.841	0.0658
Call type[a]				
exhaled moan	1.280	0.015	85.754	**< 0.0001**
inhaled grunt	-0.037	0.010	-3.836	**0.0001**
inhaled moan	0.535	0.013	40.636	**< 0.0001**
wobble	0.673	0.039	17.235	**< 0.0001**

[a] Exhale grunt was the reference level for call type

[b] Bolded values indicate $P < 0.05$

Table 4.2. Duration of the call types used in gelada male vocal sequences.

Call type	N	Mean (s)	SEM	Range (s)
Exhale grunt	2982	0.296	0.002	0.064 - 0.766
Exhale moan	300	1.597	0.430	0.770 - 4.153
Inhale grunt	872	0.244	0.004	0.029 - 0.511
Inhale moan	435	0.840	0.012	0.514 - 1.985
Wobble	41	0.948	0.127	0.261 - 4.250
Yawn	117	0.996	0.037	0.200 - 2.406

Table 4.3. Results of LMM used to test for relationships between inter-call interval duration and sequence size and interval position.

Factor	Estimate	SEM	*t*	P[a]
Intercept	0.452	0.017	26.827	**< 0.0001**
Sequence size	-0.023	0.002	-11.130	**< 0.0001**
Interval position	-0.003	0.001	-2.271	**0.0232**

[a] Bolded values indicate $P < 0.05$.

Table 4.S1. Results of LMMs used to test for relationships between call duration and sequence size and sequence position for datasets focused on each of the six call types.

Model (call type)	Factor	Estimate	SEM	*t*	P[a]
Exhale grunt	Intercept	0.306	0.007	41.655	**< 0.0001**
	Sequence size	-0.001	0.001	-1.058	0.2904
	Call position	-0.002	0.001	-2.365	**0.0181**
Exhale moan	Intercept	1.832	0.084	21.748	**< 0.0001**
	Sequence size	-0.024	0.016	-1.498	0.1353
	Call position	-0.049	0.025	-1.981	**0.0488**
Inhale grunt	Intercept	0.291	0.009	31.168	**< 0.0001**
	Sequence size	-0.010	0.001	-9.238	**< 0.0001**
	Call position	0.009	0.001	7.290	**< 0.0001**
Inhale moan	Intercept	0.806	0.030	26.483	**< 0.0001**
	Sequence size	-0.030	0.007	-4.250	**< 0.0001**
	Call position	0.040	0.008	5.232	**< 0.0001**
Wobble	Intercept	1.865	0.314	5.939	**< 0.0001**
	Sequence size	-0.075	0.037	-1.997	0.0546
	Call position	-0.012	0.035	-0.344	0.7333
Yawn	Intercept	1.00	0.088	11.380	**< 0.0001**
	Sequence size	-0.043	0.036	-1.182	0.2401
	Call position	0.053	0.036	1.459	0.1478

[a] Bolded values indicate $P < 0.05$.

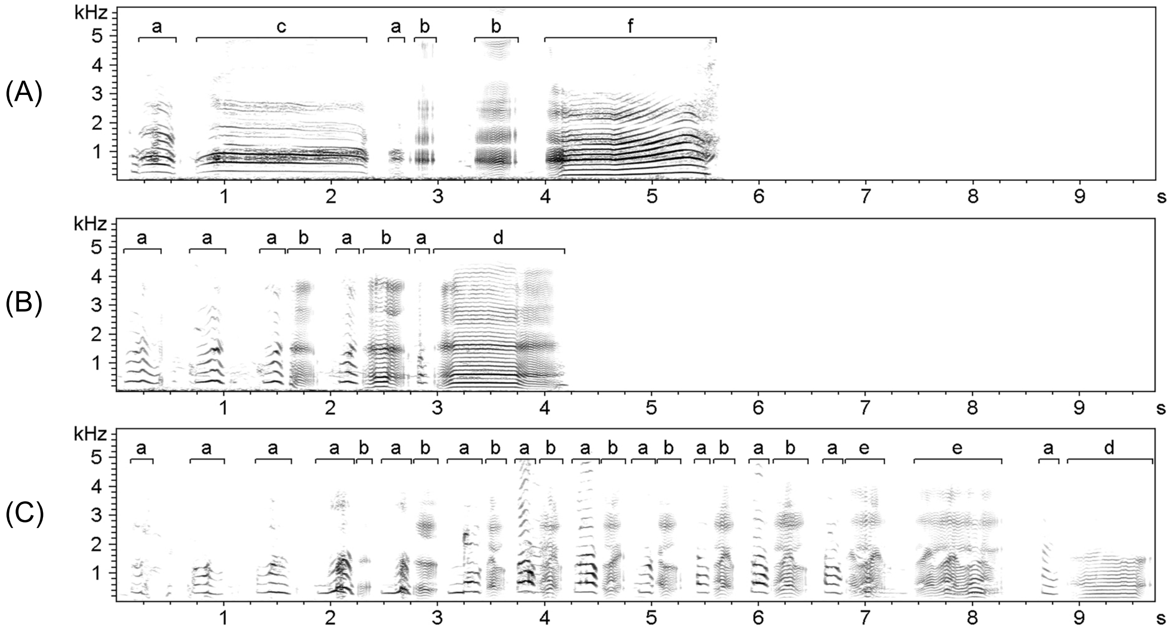


Figure 4.1. Example spectrograms of (A) a 5-call vocal sequence, (B) an 8-call sequence, and (C) a 24-call sequence from the same study male. Brackets on the spectrograms denote the start and end of each call, and the lowercase letters above the brackets represent different call types: (a) exhale grunt, (b) inhale grunt, (c) exhale moan, (d) inhale moan, (e) wobble and (f) yawn. Spectrograms were made in Avisoft (v. 5.1.12, R. Specht, Berlin).

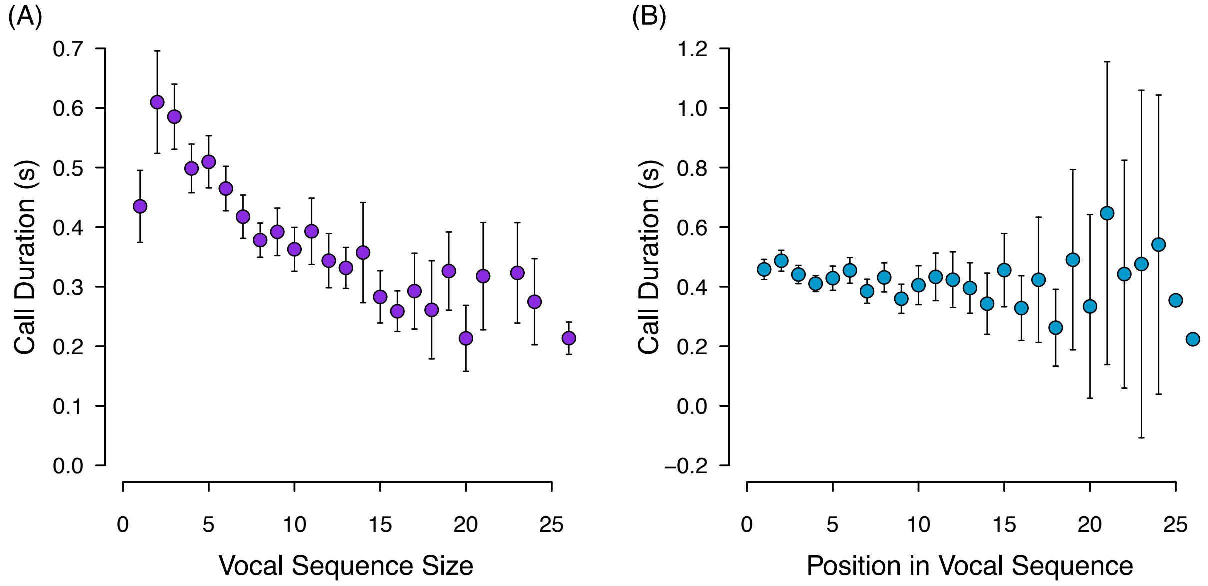


Figure 4.2. Relationship between call duration and (A) vocal sequence size (*i.e.*, number of discrete calls) and (B) call position in a vocal sequence. Points and whiskers indicate mean ± 2 SEM, and lack of whiskers indicates that there was a sample size of 1 sequence.

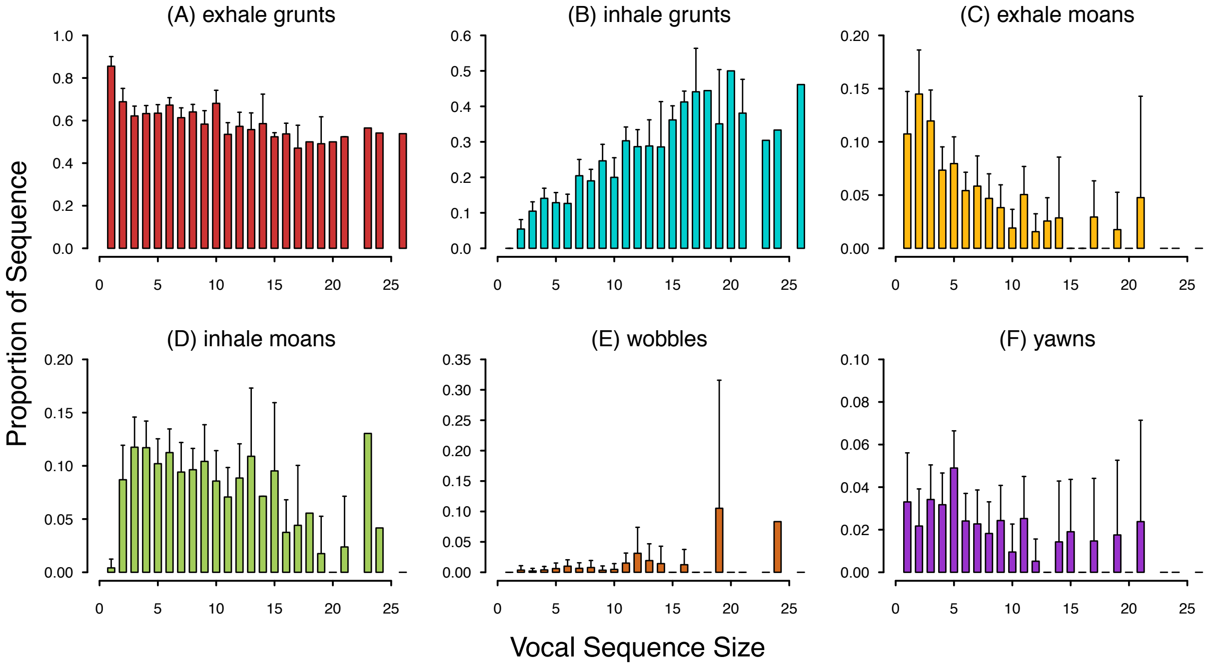


Figure 4.3. Relationship between proportions of specific call types and sequence size. Bars and whiskers indicate the mean proportion + 2 SEM of the specific call type in vocal sequences of that given size, and lack of whiskers indicates that there was no variation in the proportion values. Call types are (A) exhale grunts, (B) exhale moans, (C) inhale grunts, (D) inhale moans, (E) wobbles, and (F) yawns.

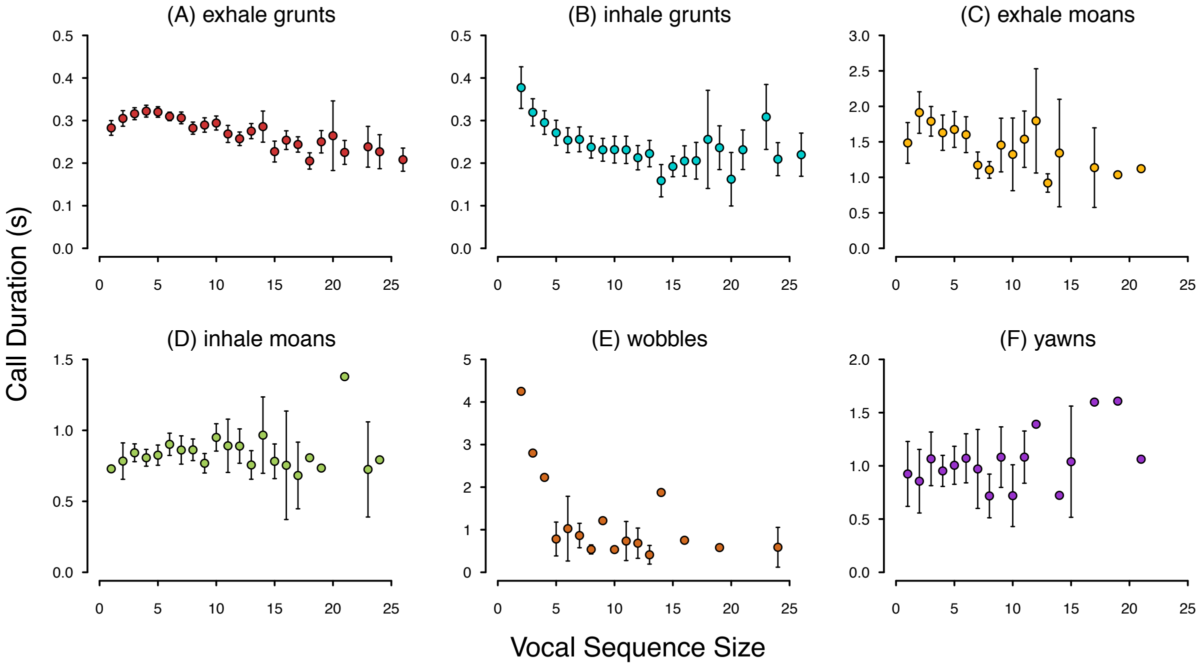


Figure 4.4. Relationship between call duration and sequence size for datasets focusing on each of the six call types. Points and whiskers indicate mean ± 2 SEM, and lack of whiskers indicates that there was a sample size of 1 sequence. Call types are (A) exhale grunts, (B) exhale moans, (C) inhale grunts, (D) inhale moans, (E) wobbles, and (F) yawns.

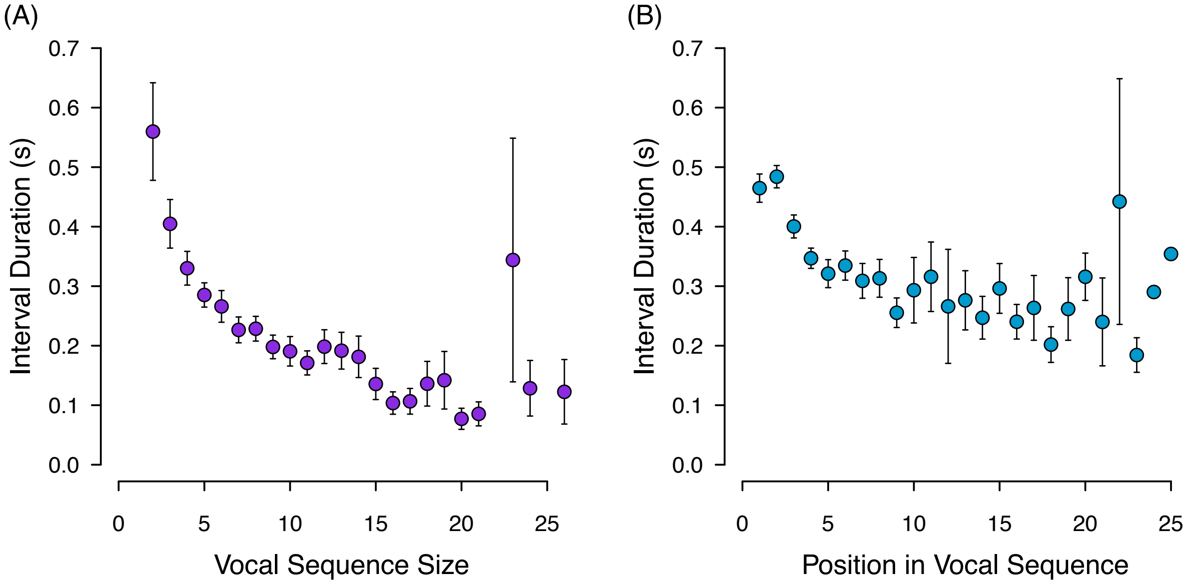


Figure 4.5. Relationship between inter-call interval duration and the (A) vocal sequence size and (B) inter-call interval position in a vocal sequence. Points and whiskers indicate mean ± 2 SEM, and lack of whiskers indicates that there was a sample size of 1 sequence.

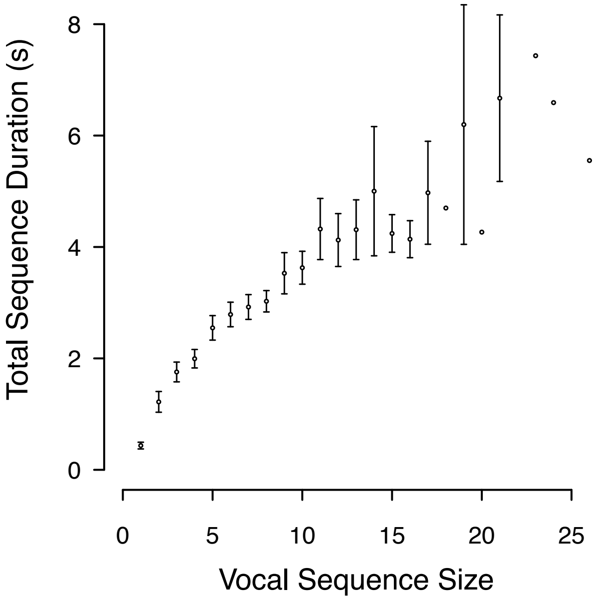


Figure 4.S1. Relationship between vocal sequence size and sequence duration (*i.e.*, total duration of sequence from start of first call to end of last call). Points and whiskers indicate mean ± 2 SEM, and lack of whiskers indicates that there was a sample size of 1 sequence.

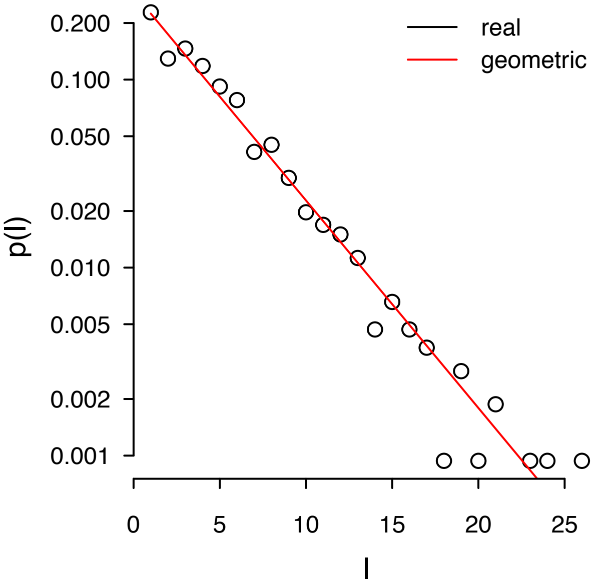


Figure 4.S2. The distribution of the vocal sequence size (black) against a geometric distribution (red) where the parameter ($\pi = 0.224$) has been obtained by maximum likelihood. $p(l)$ is the probability that a sequence consists of l calls.

Chapter 5:

Vocal complexity influences female responses to gelada male calls

Abstract

Extensive research indicates that inter-sexual selection drives the evolution of complex vocal communication in birds, but parallel lines of evidence are almost entirely absent in mammals. This dearth of evidence, particularly among primates, limits our understanding of the link between sociality and vocal complexity. Here, we use a playback experiment to quantify how wild female geladas (*Theropithecus gelada*) respond to three call types that are ‘derived’ (i.e., unique to geladas) and made by males during various affiliative contexts. These derived calls appeared to be highly salient and preferable to females: they looked longer towards and spent more time in proximity to playbacks of male vocal sequences containing one of the derived calls than to sequences containing only common and less elaborate ‘grunt’ calls. Our results provide the first experimental evidence for vocal elaboration as a male-specific strategy to maintain social bonds with females in non-human primates.

Introduction

Human's ability to combine sounds together into an endless array of meaningful words and sentences is unique, making the evolutionary roots of language a focus of intense interest[1,2]. Despite this uniqueness, several aspects of language can be studied comparatively [3,4]. In particular, there is tremendous interest in documenting the diversity of 'vocal complexity' in animals [5]. Vocal complexity is typically defined as the number of different vocalizations a species can make, or vocal repertoire size, and this trait differs extensively across taxa [6–8]. This diversity is useful because it allows for comparative studies that can identify the main types of selective pressures driving the evolution of complex forms of communication. Several comparative studies in birds and mammals, including humans, suggest an important role for social pressures in the evolution of vocal complexity by showing that broad measures of vocal complexity (e.g., repertoire size) are positively associated with sociality (e.g., group size) [6,7,9–12]. However, our understanding of the specific social functions of individual features of vocal systems, such as complex strings of sound, is more limited.

The most comprehensive data on the social functions of complex strings of sound come from research on inter-sexual selection in bird song [13–15]. Male songbirds (Passeriformes) often produce songs during courtship, and females show preference for males with larger repertoires of syllable, phrase or song types [16–21] and males producing songs composed of more complex elements [14,22,23]. In either case, sexual selection of songbird vocal complexity appears to act at the level of the "sequence", meaning that the functional unit of sound is the combination of elements rather than the individual elements themselves. While individual elements can affect responses to the song, each song element has a similar shared function (e.g., mate attraction). In some species, females may gain direct or indirect benefits by using vocal complexity as an index of male quality [24,25], while for other species, vocal complexity may primarily be a way for males to exploit females' auditory sensory biases [19]. Regardless of the specific pathway leading to a preference, the outcome is the same: in many bird species, vocal complexity can facilitate inter-sexual social interactions. However, we do not know if these findings are unique to songbirds and their unusual vocal system. The function of complexity in other animal vocal systems remains largely unexplored.

Most relevant for understanding the origins of language are the vocal systems of primates. Humans and other primates exhibit several homologies in the brain circuitry involved with communication [26], and like humans, many primates maintain long-term relationships and live in large social groups [27]. These shared traits make primates useful comparative models to explore the role that sociality plays in the evolution of vocal complexity. Yet, unlike songbirds, non-human primates exhibit relatively small repertoires of discrete and graded sounds (i.e., calls) and much less is known about how and why they combine multiple call types into sequences [4]. Broad comparative research suggest that primate species with large vocal repertoires are characterized by living in large social groups and spending a great deal of time engaged in affiliative behaviors [12]. Moreover, narrower comparative studies show that some non-human primate taxa exhibit a greater degree of vocal complexity than their close relatives because they produce more types of calls during within-group aggressive or affiliative social interactions [9,28]. Together, this body of comparative work suggests that some primates have more complex vocal systems because they produce unique 'derived' call types that evolved to better facilitate social interactions. For example, mandrills (*Mandrillus sphinx*) produce a unique suite of long distance contact calls that are thought to play a species-specific role in coordinating group movements in densely forested environments; analogous call types are not found in closely related species like baboons and geladas [29].

Identifying derived call types and the social contexts in which they are produced are the first steps towards understanding how vocal complexity functions in primates. We also need complimentary studies that examine how these derived call types are perceived by potential receivers, which we currently know little about in primates and other mammals [30–32]. We address this gap by investigating behavioral responses to derived calls in wild geladas (*Theropithecus gelada*), a primate known for its large and unique vocal repertoire and for which derived call types (i.e., calls with no clear analogs in the vocal repertoires of their close baboon relatives) have already been identified [28,33–36]. Three of these derived call types – “moans”, “wobbles”, and vocalized “yawns” – are of particular interest because they are produced almost exclusively by adult males and are the most acoustically elaborate of the derived calls. Moans are long in duration, wobbles have a high degree of frequency modulation, and yawns take up a large frequency bandwidth [28,34]. Due to the male-biased production and their elaborate form, it appears that these calls are sexually selected call types, although this possibility has never been experimentally tested.

Gelada males typically produce moans, wobbles and yawns in vocal sequences, and they do this by combining them with a homologous call type – exhaled grunts – and another ‘derived’ call type – inhaled grunts – both of which are commonly produced by both male and female geladas [36]. Males produce grunt sequences containing none, a single, or multiple elaborate derived calls during close-range affiliative social interactions (e.g., approaches, allogrooming and after female-female conflicts) with females [28,33,36]. As in birds, geladas combine different types of sounds in a single social context, suggesting that the function of each element is to add to the sequence complexity rather than to serve a unique social function. However, it remains unknown whether the vocal sequences containing elaborate derived calls and grunts elicit different responses from those containing only grunts. One intriguing possibility is that, as in birdsong, the diversification of gelada males' affiliative vocal sequences may function to attract or bond with their female counterparts [28]. Geladas aggregate into extremely large groups of over 1000 individuals that are made up of smaller 'harem-like' reproductive units composed of a dominant leader male, up to a few subordinate follower males, several females and their dependent offspring [37]. Leader males of reproductive units that use effective strategies to maintain their long-term social relationships may decrease the chance of being cuckolded by within-unit subordinate males [38,39] or out-competed by a non-unit 'bachelor' [40]. It is still unknown in geladas, and in primates more generally, whether vocal sequences containing elaborate derived calls from males influence female behavior in a way that could benefit male fitness. Such a finding would be the first evidence for inter-sexual selection of vocal complexity in a non-human primate.

We build off prior studies on the production of vocal complexity in male geladas by examining the perception of vocal complexity by female geladas. Using an experimental playback design, we modeled established behavioral assays to assess female responses to male vocal displays [41–44]. First, we tested whether or not female geladas discriminate between vocal sequences that do or do not contain one of the elaborate derived calls and grunts by comparing females' visual orientation towards playbacks of male vocal sequences. Second, we tested whether females show a 'preference' for these derived calls by comparing the amount of time that females spent in proximity to these same playbacks of derived call and grunt only sequences.

Methods

Study site and animals

Experimental playback data were collected from February to June 2014 in the Sankaber area of Simien Mountains National Park, Ethiopia. Research was approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan and was carried out in accordance with the laws and approved guidelines of Ethiopia. Study subjects were 36 adult female geladas from outside the three main study bands followed by the University of Michigan Gelada Research Project since 2006. Females from outside of the main study bands were chosen so they would be uniformly unfamiliar with the males from which we recorded playback stimuli. Although the vocalizations used here are typically exchanged between familiar individuals, it is very difficult to conduct realistic playback trials among members of the same unit as they are usually within visual contact of each other. In addition to being more tractable, using unfamiliar callers and subjects also controls for variation in relationship quality between caller and subject, ensuring that any differences in responses are likely due to differences in the stimuli. Female subjects were habituated to humans on foot up to 3-5 m and could be identified by unique body markings (e.g., ear tears and coloration).

Playback stimuli and experimental design

Vocal sequences were recorded from 12 adult male geladas who were unit leaders and/or followers from one of the three main study bands between 2008-2014. These recordings were made from less than 10 m using a Sennheiser ME-66 directional microphone and a Marantz PMD 660 or 661 digital recorder. Playback stimuli were made using PRAAT 5.2.29 for Macintosh. Each playback stimulus was composed of 2-9 calls from a natural sequence with a high signal-to-noise ratio. The majority of recorded vocal sequences were manipulated (e.g., excluding call(s) from the beginning and/or end of a sequence) to produce playback stimuli that were clear of overlapping calls and were of a similar overall duration. The amplitude of extraneous sounds (e.g. bird chirp or vocalization from another gelada) found in the intervals between calls was dampened using PRAAT.

The playback stimuli consisted of 18 ‘grunt only’ and 18 ‘derived call’ vocal sequences. Grunt only sequences were composed of exhaled grunts and inhaled grunts (mean ± SE [range]: 5.000 ± 0.406 [3 - 9] calls per stimulus). Derived call vocal sequences (3.222 ± 0.4759 [2 - 9] calls per stimulus) were composed of exhaled and inhaled grunts mixed with one of the three elaborate derived calls: exhaled moans (n = 6 stimuli; 2.500 ± 1.021 [2 - 4] calls per sequence), exhaled (n = 3) or inhaled (n = 3) wobbles (4.500 ± 1.837 [2 - 9] calls per sequence), and inhaled vocalized yawns (n = 6; 2.667 ± 1.089 [2 - 4] calls per sequence). The elaborate derived calls were acoustically different from exhaled and inhaled grunts in various ways: exhaled moans, wobbles and yawns used in the derived call playback stimuli were longer in duration compared to exhaled and inhaled grunts, wobbles had the highest frequency modulation, and yawns had the highest formant (F1) frequency (Table 5.S1). For further descriptions of these call types, see Gustison et al. [28], and see Figure 5.1 for spectrograms of grunt only and derived call sequence playback stimuli.

Six study males contributed one grunt only and one derived call sequence and six study males contributed two grunt only and two derived call sequences. Each of the grunt only sequences was paired with a derived call sequence – forming 18 playback ‘sets’. A counterbalanced matched-control design was used; 18 of the study females were presented with a grunt only sequence first and a derived call sequence second, and the other 18 females were presented with a derived call sequence first and a grunt only sequence second. Therefore, the 36 female subjects were each exposed to two playback stimuli (grunt only and derived call vocal sequences) for a total of 72 playback trials. This repeated measure design ensured that the variation in internal (e.g., reproductive state) and external (e.g., recent interactions with unit members) factors were similar for female subjects exposed to grunt only and derived call playbacks.

Several precautions were taken to ensure that playback stimuli of grunt only and derived call sequences were similar except for the call composition. First, we controlled for other acoustic signals that could affect female responses, like inter-male variation in fundamental frequency, by matching male callers across playback sets. For every playback set that consisted of a grunt only sequence from male A and a derived call sequence from male B, there was a corresponding playback set that consisted of a grunt only sequence from male B and a derived call sequence from male A. Second, variation in female response due to vocal sequence duration was controlled for by matching the durations of grunt only (2.768 s ± 0.652 [1.411-3.782]) and derived call (2.637 s ± 0.621 [1.698-3.680]) vocal sequences in a playback set (Wilcoxon signed-rank test: W = 123, N = 18, p = 0.1084). Vocal sequence duration also was similar across the 3 types of derived call sequences (Mann-Whitney U tests: moan (3.031 s ± 1.237 [2.264-3.680]) vs. wobble (2.447 s ± 0.999 [1.945-3.170]): U = 29, N1 = N2 = 6, p = 0.0931; moan vs. yawn (2.432 s ± 0.993 [1.698-3.361]): U = 28, N1 = N2 = 6, p = 0.132; wobble vs. yawn: U = 20, N1 = N2 = 6, p = 0.8182). Third, grunt only and derived call sequences were played at a similar decibel level. A Radioshack Digital Sound Level Meter was used to check that the maximum dB of each sequence in a quiet indoor environment was similar for grunt only (69.833 dB ± 16.460 [67-75]) and derived call sequences (69.667 dB ± 16.421 [64-78]; Wilcoxon signed-rank test: W = 57.5, N = 18, p = 0.7764). The maximum dB also was similar across types of derived call sequences (Mann-Whitney U tests: moan (68.167 dB ± 27.829 [64-72]) vs. wobble (70.667 dB ± 28.850 [67-78]): U = 11.5, N1 = N2 = 6, p = 0.3315; moan vs. yawn (70.167 dB ± 28.645 [67-74]): U = 10.5, N1 = N2 = 6, p = 0.2573; wobble vs. yawn: U = 15.5, N1 = N2 = 6, p = 0.7462).

Playback protocol

An adult female was chosen as a subject for a playback trial if she was relatively stationary (i.e., feeding or resting), was not engaged in social activity, and was close to vegetation where the speaker could be hidden. No individuals were located between the subject and the speaker. During playback sessions, playback stimuli were presented with a SanDisk Clip mp3 player connected to a Bose Roommate II portable loudspeaker. The speaker was concealed behind vegetation 3-10 m from the subject in the direction of the gelada herd. Thus, playback stimuli presumably represented vocal sequences from unit males rather than bachelors [47]. Female behavior was recorded with a Kodak PlaySport HD waterproof pocket video camera. The playback stimulus was played after 10 s of video recording if the subject remained engaged in non-social stationary behavior and her body and head were oriented 90 degrees from the speaker. The subject continued to be video recorded for one minute following the playback stimulus. Previous research has shown that one minute adequately captures gelada responses to contact calls such as grunts [47]. The second trial involving the same subject was played at least two minutes after the end of recording the first trial from a location that was at least 3 m away from the location of the first playback stimulus. At the end of each playback set, the relative locations and distances between the video recorder, closest adult unit male, speaker and study female (at both the presentation of the playback stimuli and after one minute) were recorded. The presentation order of different playback sets was randomized across subjects.

Following the guidelines put forth by Fischer et al. [43], several precautions were made to avoid habituation of female subjects to the playback stimuli. For example, we played male vocal sequences at a much lower rate than they occur naturally; we played no more than three playback sets per day (6 vocal sequences), whereas gelada males naturally produce vocal sequences including grunts and derived calls at least 14 times per hour on average [28]. Given that geladas range in herds that include many males, females are exposed to these vocal sequences at a much higher rate [28]. For playback sets carried out on the same day, we chose female subjects from different units that were out of sight (and likely audible range) from the location of the previous playback set. Playback stimuli were never repeated on the same day.

Analysis of female behavioral responses to playback stimuli

An independent observer scored behavior using Windows Live Movie Maker 2011 (Microsoft, Redmond, WA, USA) from at least 10 seconds before to over one minute after the playback stimulus using frame-by-frame analysis (behavioral responses were later calculated from the time frame from the end of the playback stimuli to one minute after). The sound was turned off so that the observer was blind to when and what type of playback stimulus was played. The observer scored the start and end of all visual orientation towards the speaker, defined as the subject's head being oriented ± 45 degrees in the direction of the speaker. From these data, we computed three specific visual orientation measures: duration of the first visual orientation towards the speaker, duration of the total visual orientation towards the speaker, and the number of separate visual orientations towards the speaker. All behavioral measures were taken from the end of the playback sequence so that it could be assured that we were quantifying the female responses to an entire sequence. Because of this, we do not measure a females latency to look, which is common in playback experiments with non-human primates [43]. Females often looked towards the speaker prior to the conclusion of the sequence, and so lag to look is not a relevant measure of female response to the entire playback stimulus. The independent observer also scored the videos for the amount of time that the female spent in proximity to the speaker, defined as the time post-stimulus until the female moved over 1 m in a direction 90-180 degrees from the speaker. We focus on proximity behavior rather than other potential ‘preference’ behaviors such as approach or copulation displays because the social structure of wild geladas is such that females do not commonly engage in close-range affiliative interactions (e.g., approach, grooming, and copulation) with non-unit males. Previous studies suggest that female geladas do not mate with non-unit males, and if they mate with subordinate follower males in their units, they are at risk of aggression from the leader male [38,39]. In the present study, there was only one instance following a grunt only sequence that the female subject moved closer to the speaker.

To check for intra-observer reliability, the observer re-scored each video for a second time at least 56 days after the original scoring. Rho values from Spearman signed-rank correlations were used to assess consistency in the four analyzed behaviors. All behaviors had intra-observer reliability rho values over 0.84 (duration of first visual orientation – 0.877; duration of total visual orientation – 0.876; number of looks – 0.863; time spent in proximity to speaker – 0.840).

We used Shapiro-Wilk tests to check whether behavioral variables fit a normal distribution. No variables fit a normal distribution and this remained true after transformation (square-root transformation, $p < 0.05$; natural log transformation, $p < 0.05$). Therefore, we used non-parametric tests. For each behavioral measure, we first used Wilcoxon sign-rank tests to see if females responded differentially to grunt only and derived call vocal sequences. We did not compare responses to different types of derived calls because of small sample sizes. Second, we used Wilcoxon sign-rank tests to see if there were any order effects by comparing within-female responses to the first and second playback stimulus. All tests were two-tailed (**α**= 0.05) and carried out using Cran R package version 3.1.1. Descriptive statistics are reported as mean ± SE [range].

Results

Visual orientation towards the speaker

There was good evidence that female geladas distinguish between playbacks of grunt only and derived call sequences from unfamiliar males (Figure 5.1). The first visual orientation that females made towards the speaker in the minute following the conclusion of playback stimuli was longer following derived call sequences (mean ± SE [range]: 1.139 s ± 0.190 [0.000 – 4.370 s]) than to grunt only sequences (0.421 s ± 0.070 [0.000 – 1.980 s]; W = 73.5, N = 36, p = 0.0007; Figure 5.2A). Females also spent more time overall visually orienting towards playbacks of derived call sequences (1.834 s ± 0.305 [0.000 – 7.970 s]) than to grunt only sequences (0.873 s ± 0.145 [0.000 – 4.510 s]; W = 134, N = 36, p = 0.0261; Figure 5.2B). There was no evidence that females made different numbers of separate visual orientations towards the speaker following playbacks of derived call (1.583 looks ± 0.264 [0 – 7 looks]) and grunt only sequences (1.417 looks ± 0.236 [0 – 6 looks]; W = 113, N = 36, p = 0.6669; Figure 5.2C). Visual orientation towards the three types of derived call playbacks were similar, although on average, females tended to looked longer towards sequences including wobbles or yawns than those including moans (Table 5.1).

There was no evidence that the presentation order of playback stimuli affected females' visual orientation responses to vocal sequences from unfamiliar males. The first visual orientation that females made towards the speaker in the minute following the playback stimuli was similar following the vocal sequence of the first playback trial (0.913 s ± 0.152 [0.000 – 4.089 s]) compared to the second playback trial (0.648 s ± 0.108 [0.000 – 4.370 s; W = 322.5, N = 36, p = 0.1470]). Females also spent a similar amount of time overall visually orienting towards the vocal sequence of the first playback trial (1.524 s ± 0.254 [0.000 – 6.482 s]) as to the second playback trial (1.183 s ± 0.197 [0.000 – 7.970 s]; W = 305, N = 36, p = 0.2682). Additionally, females made a similar number of distinct visual orientations (i.e., "looks") towards the speaker following the vocal sequence of the first playback trial (1.639 looks ± 0.273 [0 – 7 looks]) as to the second playback trial (1.361 looks ± 0.227 [0 – 4 looks]; W = 163.5, N = 36, p = 0.2270).

Time spent in proximity to the speaker

There was evidence to indicate that females spent more time in proximity to the speaker following playbacks of derived call sequences (58.775 s ± 9.796 [17.027 – 60.000 s]) than to grunt only sequences (48.836 s ± 8.139 [1.501 – 60.000 s]; W = 0, N = 36, p = 0.0039; Figure 5.3). Proximity behavior was the same towards the three types of derived call playbacks in that females almost always spent at least one minute in close proximity to the speaker upon hearing a moan, wobble or yawn sequence (Table 5.1). There was no evidence that the presentation order of playback stimuli affected the amount of time that females spent in proximity to the speaker. Females spent a similar amount of time in proximity to the speaker following the vocal sequence of the first playback trial (56.139 s ± 9.357 [11.820 – 60.000 s]) as to the second playback trial (51.472 s ± 8.579 [1.501 – 60.000 s]; W = 50, N = 36, p = 0.1424).

Discussion

This is the first study to systematically show that vocal complexity may be driven by inter-sexual selection in a non-human primate. We found that female geladas clearly distinguish between derived and homologous calls and the direction of the differences in responding all suggest a stronger salience of and, possibly a preference for, the derived calls. Specifically, females hearing playbacks of male vocal sequences containing one of three derived call types – moans, wobbles and yawns – oriented longer as well as spent more time in proximity to the speaker. These results align with an extensive body of research in songbirds demonstrating that diversified male vocal signals may function to attract mates and maintain long-term social bonds [16–21].

Although similar evidence in mammalian species is relatively sparse [30], our results do support a growing body of work in rodents and bats suggesting that vocal complexity may have analogous social functions in mammals. As with female geladas, female mice (*Mus musculus*) spend more time around playbacks of male song containing many elaborate syllable types than those composed of only a simple syllable type [41]. In addition, male greater sac-winged bats (*Saccopteryx bilineata*) producing songs with several unique syllable types have more females who consistently roost in their harem territories than do males producing fewer syllable types [45]. Therefore, both male geladas and sac-winged bats utilizing a a more complex string of sounds may be better equipped to maintain bonds with the females in their harem-like reproductive units. Thus, these results are the first to indicate that the diversification of call types may have evolved as a male-specific strategy to maintain long-term social bonds with females in primates.

Our findings also have implications for gelada society and the evolution of tolerance. Females had generally weak responses and stayed in proximity to derived call sequences (and many of the grunt sequences) even though the males vocalizing were unfamiliar to the subject. This is surprising considering that the close and sudden presence of a stranger is a rare and potentially distressing event for primates that tend to live in long-term and relatively stable social groups [46]. One likely correlate of this apparent tolerance of strangers is that females may simply be unable to recognize the vocal signals of males from their band but outside of their reproductive units and are consequently quite habituated to hearing calls from unrecognized individuals. This corresponds with previous evidence showing that male geladas do not distinguish vocal sequences of familiar males from unfamiliar males [47]. Another possibility is that, in addition to a primary function of derived calls in male-female bonding within units, the calls may have an inherent attractiveness that leads to a secondary function in maintaining cohesion across units. Geladas live in fission-fusion societies, and it is a regular occurrence for reproductive units to travel with unfamiliar units [37]. Derived calls may play a particularly important role in coordinating these flexible group dynamics. Additionally, female composition in the reproductive units is stable, but leader male tenures rarely last more than a few years [39]. This means that females need to rapidly form strong associations with new leader males following takeovers. An intriguing line of future research will be to explore how new leader males may use derived calls as a strategy to develop their social bonds with females.

The main caveat of the present study is that it is still unknown whether females are attending specifically to ‘derived calls’ or to ‘complex vocal sequences’. Orientation responses are notoriously difficult to interpret [43], and so derived calls may invoke a greater orientation response than grunts because they are rarer or more indicative of salient social stimuli instead of reflecting female interest [28]. However, females stayed close to playbacks of derived calls, suggesting that it is not simply the case that those calls are startling. Also, given that geladas live in large fission-fusion societies and vocalize at a high rate [28], it is not unusual for females to hear complex vocal sequences from unfamiliar males. Instead, our findings indicate that females may show a preference for innovative vocal signals, which would align with studies of non-primate taxa like zebra finches (*Taeniopygia guttata*) [19]. Derived calls also are characterized by acoustic properties that may make them more elaborate and potentially more attractive than the typical grunt (e.g., long duration, frequency modulation and large bandwidth). Such acoustic properties are also found in call types preferred by female birds [22,23,48,49], anurans [50], and other mammals [41,45,51]. Conversely, female geladas may be attending to the degree of complexity in vocal sequences rather than to the specific use of derived calls. This explanation would align with studies showing that female birds are attracted to vocal stimuli containing large syllable and song repertoires [16–21]. Even in many of these bird studies, however, it is unclear whether females are exhibiting a preference for large repertoires or whether repertoire diversification is actually driven by a greater signal value in individual components [14]. In both birds and geladas, it is difficult to untangle these explanations because presenting females with individual components alone rather than in sequences would be an unnatural stimulus that may evoke responses that are not ecologically relevant.

Future studies are needed to tease apart the types of information gained from different vocal sequence combinations made by geladas. Females in our study did not appear to distinguish between sequences containing different derived call types, but small sample sizes prevented us from making statistical comparisons. One exciting possibility is that derived call vocal sequences are honest indicators of mate quality (being produced at a higher rate by the best quality males), which would make them particularly attractive to females [52]. Given that every male used in this study contributed equal numbers of derived call and grunt sequences, it is unlikely that our results reflect female preference for the acoustic qualities of specific males [42]. Instead, it seems more likely that gelada females prefer males giving a higher output of derived call vocal sequences. Further work will be needed to test the possibility that variation in the complexity of male vocal behavior translates to reproductive success. Until then, we are unable to completely rule out the possibility that variation in female orientation and proximity behavior in response to male calls reflect differences in motivation to engage with a *social* partner rather than a *sexual* partner. Additionally, these derived call sequences may serve an alternative or complimentary role in female detection and identification of unit males [53]. In other words, derived call sequences could counteract environmental noise such as the chorus of other geladas in their large communities. Similar solutions to cope with conspecific noise has been proposed for other species such as Túngara frogs (*Physalaemus pustulosus*) [50].

There is great debate over the evolutionary origins of highly complex and diversified forms of communication such as human language. One focus has been on investigating comparative evidence for semantic communication as a key driver in the evolution of complex communication [54,55]. Despite the small vocal repertoires of non-human primates, there are many well-studied examples of how diversified primate calls and call combinations may have evolved as a tool to communicate functionally referential information about food and predators [56–58]. Another focus has been on investigating comparative evidence for elaboration in affective communication, or types of holistic communication that may seem multi-faceted in appearance but serve singular social functions [59]. While best exemplified by the elaborate bird songs that function to successfully interact with mates and deter rivals [15], there is only limited evidence that non-human primate vocal systems can be elaborate in similar contexts (e.g., gibbons (*Hylobates sp.*) [60]). By providing novel evidence of a potential role for sexual selection in the emergence of vocal complexity in primate vocalizations, our results build support for non-semantic vocal elaboration as an early step towards language [34].

Acknowledgements

MLG was supported by grants from the National Geographic Society (Gr# 9122-12, W304-14), Leakey Foundation, University of Michigan, and Petridish (primary supporter: John Allen). TJB and the gelada project were supported by the Wildlife Conservation Society (SSF Gr# 67250), the National Geographic Society (Gr# 8100-06), the Leakey Foundation, the National Science Foundation (Gr# BCS-0715179, BCS-0962118), and the University of Michigan. We are grateful to the Ethiopian Wildlife

Conservation Authority and the wardens and staff of the Simien Mountains National Park for granting us permission to conduct this research in Ethiopia. Marcela Benitez, Megan Gomery, Elizabeth Johnson, Levi Morris, Ambaye Fanta and Tara Regan for helped with conducting experiments, Aliza le Roux provided access to acoustic recordings, and Niki Abdou helped with video scoring.

References

1. Bickerton, D. Language evolution: A brief guide for linguists. *Lingua* **117,** 510–526 (2007).
2. Pinker, S. & Jackendoff, R. The faculty of language: what's special about it? *Cognition* **95,** 201–36 (2005).
3. Berwick, R. C., Okanoya, K., Beckers, G. J. L. & Bolhuis, J. J. Songs to syntax: The linguistics of birdsong. *Trends Cogn. Sci.* **15,** 113–121 (2011).
4. Fedurek, P. & Slocombe, K. E. Primate vocal communication: a useful tool for understanding human speech and language evolution? *Hum. Biol.* **83,** 153–73 (2011).
5. Freeberg, T. M., Ord, T. J. & Dunbar, R. I. M. The social network and communicative complexity: preface to theme issue. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367,** 1782–1784 (2012).
6. Blumstein, D. T. & Armitage, K. B. Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *Am. Nat.* **150,** 179–200 (1997).
7. Manser, M. B. *et al.* Vocal complexity in meerkats and other mongoose species. *Adv. Study Behav.* **46,** 281–310 (2014).
8. Wilkinson, G. S. in *Anim. Soc. Complex. Intell. Cult. Individid. Soc.* (eds. de Waal, F. B. M. & Tyack, P. L.) 322–341 (Harvard University Press, 2003).
9. Bouchet, H., Blois-Heulin, C. & Lemasson, A. Social complexity parallels vocal complexity: a comparison of three non-human primate species. *Front. Psychol.* **4,** 390 (2013).
10. Freeberg, T. M. Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychol. Sci.* **17,** 557–61 (2006).
11. Hay, J. & Bauer, L. Phoneme inventory size and population size. *Language (Baltim).* **83,** 388–400 (2007).
12. McComb, K. & Semple, S. Coevolution of vocal communication and sociality in primates. *Biol. Lett.* **1,** 381–5 (2005).
13. Catchpole, C. K. Bird song, sexual selection and female choice. *Trends Ecol.*

Evol. **2,** 94–97 (1987).

14. Byers, B. E. & Kroodsma, D. E. Female mate choice and songbird song repertoires. *Anim. Behav.* **77,** 13–22 (2009).
15. Nowicki, S. & Searcy, W. A. Song and mate choice in birds: How the development of behavior helps us understand function. *Auk* **122,** 1–14 (2005).
16. Baker, M. C., Bjerke, T. K., Lampe, H. U. & Espmark, Y. O. Sexual response of female yellowhammers to differences in regional song dialects and repertoire sizes. *Anim. Behav.* **35,** 395–401 (1987).
17. Buchanan, K. L. & Catchpole, C. K. Female choice in the sedge warbler Acrocephalus schoenobaenus: multiple cues from song and territory quality. *Proc. R. Soc. B Biol. Sci.* **264,** 521–526 (1997).
18. Catchpole, C. K. Sexual selection and the evolution of complex songs among European warblers of the genus *Acrocephalus*. *Behaviour* **74,** 149–165 (1980).
19. Collins, S. A. Is female preference for male repertoires due to sensory bias? *Proc. R. Soc. B Biol. Sci.* **266,** 2309–2314 (1999).
20. Mountjoy, D. J. & Lemon, R. E. Female choice for complex song in the European starling: a field experiment. *Behav. Ecol. Sociobiol.* **38,** 65–71 (1996).
21. Reid, J. M. *et al.* Song repertoire size predicts initial mating success in male song sparrows, *Melospiza melodia*. *Anim. Behav.* **68,** 1055–1063 (2004).
22. Vallet, E. & Kreutzer, M. Female canaries are sexually responsive to special song phrases. *Anim. Behav.* **49,** 1603–1610 (1995).
23. Rehsteiner, U., Geisser, H. & Reyer, H. Singing and mating success in water pipits: one specific song element makes all the difference. *Anim. Behav.* **55,** 1471–81 (1998).
24. Darolová, A., Krištofík, J., Hoi, H. & Wink, M. Song complexity in male marsh warblers: does it reflect male quality? *J. Ornithol.* **153,** 431–439 (2012).
25. Hiebert, S. M., Stoddard, P. K. & Arcese, P. Repertoire size, territory acquisition and reproductive success in the song sparrow. *Anim. Behav.* **37,** 266–273 (1989).
26. Ghazanfar, A. A. & Eliades, S. J. The neurobiology of primate vocal communication. *Curr. Opin. Neurobiol.* **28,** 128–135 (2014).
27. Aiello, L. C. & Dunbar, R. I. M. Neocortex size, group size, and the evolution of

language. *Curr. Anthropol.* **34,** 184–193 (1993).

28. Gustison, M. L., le Roux, A. & Bergman, T. J. Derived vocalizations of geladas (Theropithecus gelada) and the evolution of vocal complexity in primates. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **367,** 1847–59 (2012).
29. Kudo, H. The study of vocal communication of wild mandrills in Cameroon in relation to their social structure. *Primates* **28,** 289–308 (1987).
30. Charlton, B. D. Experimental tests of mate choice in nonhuman mammals: the need for an integrative approach. *J. Exp. Biol.* **216,** 1127–30 (2013).
31. Slocombe, K. E., Waller, B. M. & Liebal, K. The language void: The need for multimodality in primate communication research. *Anim. Behav.* **81,** 919–924 (2011).
32. Ramsier, M. A., Cunningham, A. J., Finneran, J. J. & Dominy, N. J. Social drive and the evolution of primate hearing. *Philos. Trans. R. Soc. B Biol. Sci.* **367,** 1860–1868 (2012).
33. Aich, H., Moos-Heilen, R. & Zimmermann, E. Vocalizations of adult gelada baboons (*Theropithecus gelada*): Acoustic structure and behavioural context. *Folia Primatol.* **55,** 109–132 (1990).
34. Bergman, T. J. Speech-like vocalized lip-smacking in geladas. *Curr. Biol.* **23,** R268–9 (2013).
35. Richman, B. Some vocal distinctive features used by gelada monkeys. *J. Acoust. Soc. Am.* **60,** 718–24 (1976).
36. Richman, B. Rhythm and melody in gelada vocal exchanges. *Primates* **28,** 199–223 (1987).
37. Snyder-Mackler, N., Beehner, J. C. & Bergman, T. J. Defining higher levels in the multilevel societies of geladas (*Theropithecus gelada*). *Int. J. Primatol.* **33,** 1054–1068 (2012).
38. le Roux, A., Snyder-Mackler, N., Roberts, E. K., Beehner, J. C. & Bergman, T. J. Evidence for tactical concealment in a wild primate. *Nat. Commun.* **4,** 1462 (2013).
39. Snyder-Mackler, N., Alberts, S. C. & Bergman, T. J. Concessions of an alpha male? Cooperative defence and shared reproduction in multi-male primate

groups. *Proc. R. Soc. B Biol. Sci.* **279,** 3788–3795 (2012).

40. Pappano, D. J., Snyder-Mackler, N., Bergman, T. J. & Beehner, J. C. Social 'predators' within a multilevel primate society. *Anim. Behav.* **84,** 653–658 (2012).
41. Chabout, J., Sarkar, A., Dunson, D. B. & Jarvis, E. D. Male mice song syntax depends on social contexts and influences female preferences. *Front. Behav. Neurosci.* **9,** 1–16 (2015).
42. Charlton, B. D., Reby, D. & McComb, K. Female red deer prefer the roars of larger males. *Biol. Lett.* **3,** 382–385 (2007).
43. Fischer, J., Noser, R. & Hammerschmidt, K. Bioacoustic field research: a primer to acoustic analyses and playback experiments with primates. *Am. J. Primatol.* **75,** 643–63 (2013).
44. Tomaszycki, M. L. & Adkins-Regan, E. Experimental alteration of male song quality and output affects female mate choice and pair bond formation in zebra finches. *Anim. Behav.* **70,** 785–794 (2005).
45. Davidson, S. M. & Wilkinson, G. S. Function of male song in the greater white-lined bat, *Saccopteryx bilineata*. *Anim. Behav.* **67,** 883–891 (2004).
46. Harris, T. R. Testing mate, resource and infant defence functions of intergroup aggression in non-human primates: issues and methodology. *Behaviour* **144,** 1521–1535 (2007).
47. Bergman, T. J. Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proc. R. Soc. B Biol. Sci.* **277,** 3045–53 (2010).
48. Ballentine, B., Hyman, J. & Nowicki, S. Vocal performance influences female response to male bird song: An experimental test. *Behav. Ecol.* **15,** 163–168 (2004).
49. Drăgănoiu, T. I., Nagle, L. & Kreutzer, M. Directional female preference for an exaggerated male trait in canary (*Serinus canaria*) song. *Proc. R. Soc. B Biol. Sci.* **269,** 2525–31 (2002).
50. Rand, A. S. & Ryan, M. J. The adaptive significance of a complex vocal repertoire in a neotropical frog. *Z. Tierpsychol.* **57,** 209–214 (1981).
51. Behr, O. & von Helversen, O. Bat serenades—complex courtship songs of the

sac-winged bat (*Saccopteryx bilineata*). *Behav. Ecol. Sociobiol.* **56,** 106–115 (2004).

52. Gil, D. & Gahr, M. The honesty of bird song : multiple constraints for multiple traits. *Trends Ecol. Evol.* **17,** 133–141 (2002).
53. Endler, J. A. & Basolo, A. L. Sensory ecology, receiver biases and sexual selection. *Trends Ecol. Evol.* **13,** 415–420 (1998).
54. Seyfarth, R. M. & Cheney, D. L. Meaning and emotion in animal vocalizations. *Ann. N. Y. Acad. Sci.* **1000,** 32–55 (2006).
55. Zuberbühler, K. Referential signaling in non-human primates : Cognitive precursors and limitations for the evolution of language. *Adv. Study Behav.* **33,** 265–307 (2003).
56. Cäsar, C., Zuberbühler, K., Young, R. J. & Byrne, R. W. Titi monkey call sequences vary with predator location and type. *Biol. Lett.* **9,** 20130535 (2013).
57. Clay, Z. & Zuberbühler, K. Bonobos extract meaning from call sequences. *PLoS One* **6,** e18786 (2011).
58. Seyfarth, R. M., Cheney, D. L. & Marler, P. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science (80-.).* **210,** 801–803 (1980).
59. Locke, J. L. Cost and complexity: selection for speech and language. *J. Theor. Biol.* **251,** 640–52 (2008).
60. Geissmann, T. & Orgeldinger, M. The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Anim. Behav.* **60,** 805–809 (2000).

Table 5.1. Results for visual orientation and proximity responses to the three types of derived call playback stimuli.

Dependent variable	N trials	Mean ± SE (s)	Range (s)
First visual orientation			
Moan	12	0.905 ± 0.261	0.000 – 2.270
Wobble	12	1.333 ± 0.385	0.000 – 4.370
Yawn	12	1.179 ± 0.340	0.000 – 3.730
Total visual orientation			
Moan	12	1.172 ± 0.338	0.000 – 4.190
Wobble	12	2.394 ± 0.691	0.000 – 6.482
Yawn	12	1.937 ± 0.559	0.000 – 7.970
Number of visual orientations			
Moan	12	1.250 ± 0.361	0 – 4
Wobble	12	2.167 ± 0.625	0 – 7
Yawn	12	1.333 ± 0.385	0 – 3
Time spent in proximity to speaker			
Moan	12	60.000[a]	60.000[a]
Wobble	12	60.000[a]	60.000[a]
Yawn	12	56.326 ± 16.260	17.027 – 60.000

[a] All subjects engaged in this behavior for at least 60 seconds.

Table 5.S1. Acoustic parameters for call types used in playback stimuli, reported in mean ± SE [range].

Call type	N	Duration (s)[a]	F1 Frequency (Hz)[b]	F1 Bandwidth (Hz)	F1 Modulation (CV)
Exhaled grunt	94	0.324 ± 0.012 [0.136 – 0.766]	714.168 ± 8.084 [515.224 – 859.491]	421.950 ± 27.350 [58.464 – 1202.765]	12.988 ± 0.798 [2.336 – 34.448]
Inhaled grunt	30	0.470 ± 0.068 [0.101 – 1.430]	764.651 ± 24.096 [523.135 – 968.694]	586.732 ± 67.626 [82.359 – 1414.826]	16.350 ± 2.011 [2.592 – 47.127]
Moan	6	2.131 ± 0.274 [1.265 – 3.247]	746.734 ± 37.475 [593.810 – 857.523]	417.137 ± 69.587 [149.807 – 623.556]	9.019 ± 1.066 [5.757 – 12.319]
Wobble	6	1.129 ± 0.192 [0.507 – 1.680]	750.921 ± 38.804 [631.204 – 871.575]	707.137 ± 193.463 [227.722 – 1504.683]	19.822 ± 6.110 [7.709 – 44.466]
Yawn	6	1.321 ± 0.166 [0.728 – 1.840]	817.020 ± 29.881 [710.988 – 886.753]	619.122 ± 100.842 [333.326 – 987.294]	14.900 ± 2.438 [8.046 – 22.304]

[a] Duration calculated with Avisoft SASLab Pro.

[b] The 1st formant (F1) was calculated with PRATT using standard settings. Coefficients of variation (CV) were calculated from the listing of all first formant (F1) measures (every 0.00625 s increment) for individual calls. Fundamental frequency (F0) was not included because it could not be determined for all calls.

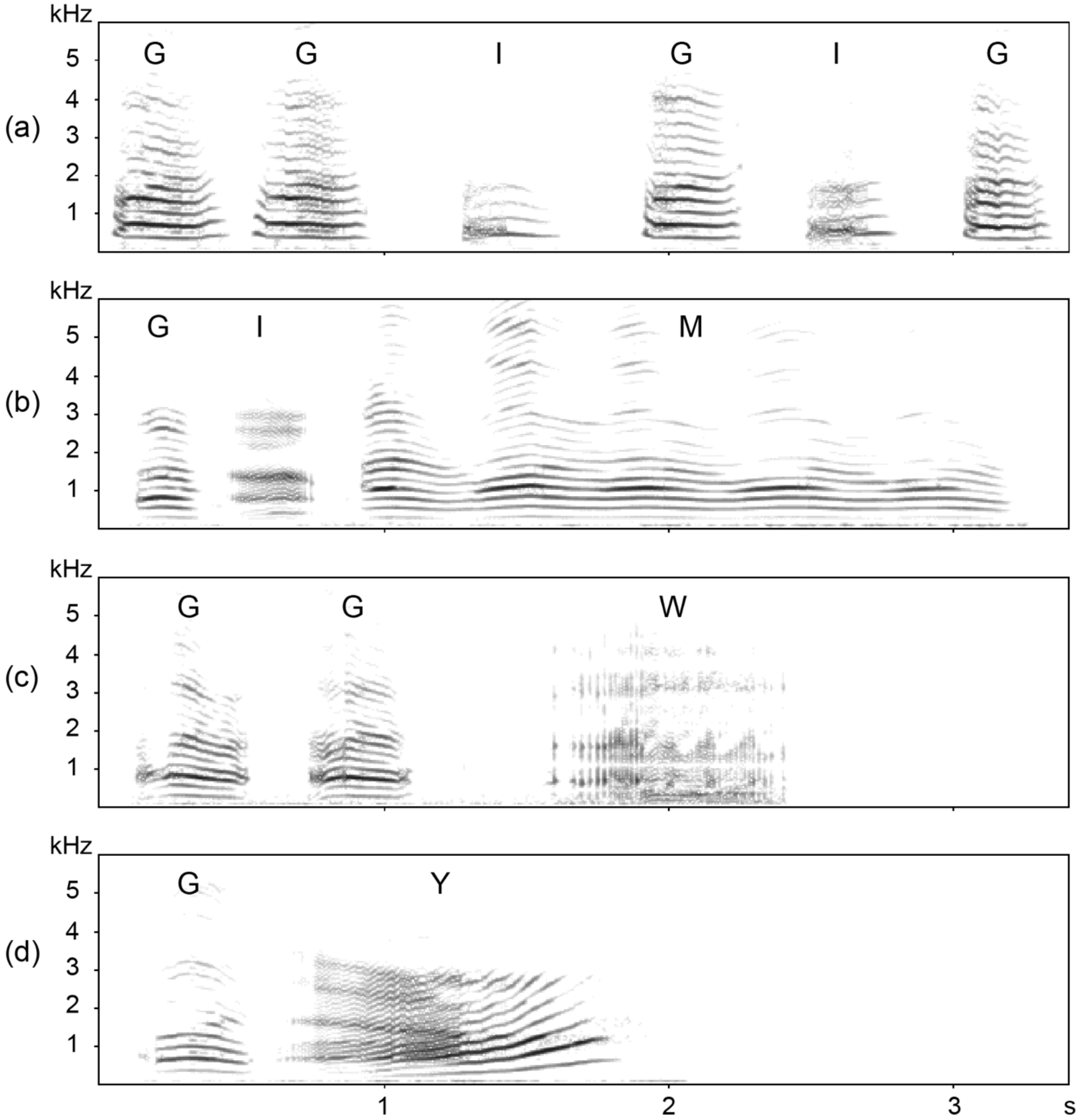


Figure 5.1. Example playback stimuli of a (a) grunt-only vocal sequence and derived call vocal sequences that include either a (b) moan, (c) wobble or (d) yawn. 'G' refers to exhaled grunts, 'I' refers to inhaled grunts, 'M' refers to a moan, 'W' refers to a wobble, and 'Y' refers to a yawn. All of the vocal sequences include exhaled and inhaled call types. Spectrograms were made with Avisoft SAS Lab Pro.

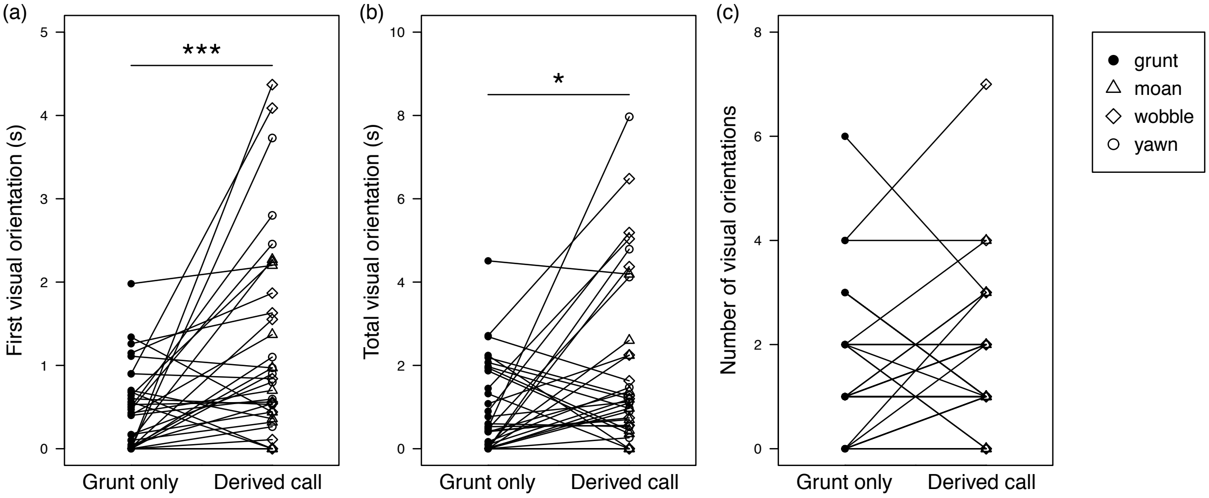


Figure 5.2. Visual orientation towards the speaker following the conclusion of unfamiliar male vocal sequences. Vocal sequences were composed of only grunts (grunt) or included one of the derived calls (moan, wobble, yawn). Behaviors measured include (a) the duration of the first visual orientation towards the speaker, (b) the duration of the total visual orientation towards the speaker, and (c) the number of distinct visual orientations towards the speaker. Lines connect trials carried out with the same female subject. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

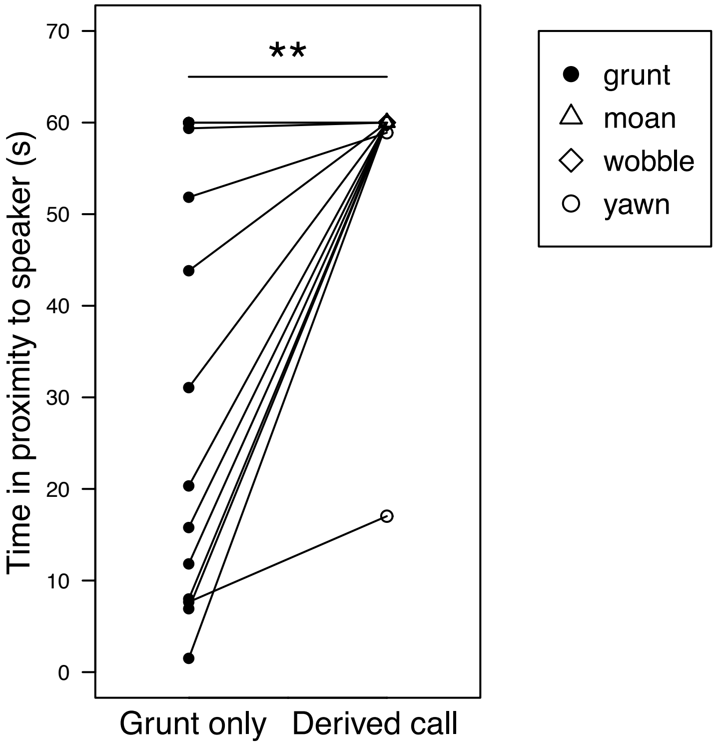


Figure 5.3. Time spent in proximity to the speaker following the conclusion of unfamiliar male vocal sequences. Vocal sequences were composed of only grunts (grunt) or included one of the derived calls (moan, wobble, yawn). Lines connect trials carried out with the same female subject. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Chapter 6:

The social functions of vocal complexity in geladas

Abstract

In support of the social complexity hypothesis, we know that more social taxa (species with larger group size or that spend more time spent socializing) tend to produce more complex vocalizations. We know very little about the specific function that vocal complexity plays within a highly social setting. Here, I address the function of vocal complexity in male geladas living in the Simien Mountains National Park in Ethiopia. I tested whether complex vocal sequences function to counteract challenges that gelada males face by living in a large group (more conspecific noise to compete with) or to help maintain their social bonds with females. I found support for both hypotheses. Vocal sequences were more complex (more derived call types) when the conspecific noise level was high. Furthermore, producing sequences with more derived calls leads to being groomed by females. Together, these findings suggest that the need to maintain cross-sex bonds within a noisy backdrop of conspecific vocalizations may favor greater vocal complexity.

Introduction

The complexity of animal vocal communication varies enormously across taxa, from humans with an infinite repertoire of sound combinations to a mongoose species that only produces three different sound types (Baker, 1988). One of the most promising evolutionary explanations for why vocalizations become more diversified comes from the social complexity hypothesis, which states that high levels of sociality (e.g., group size, amount of time socializing) led to the evolution of more complex forms of communication (Blumstein & Armitage, 1997; Freeberg, 2006; Wilkinson, 2003). To date, this hypothesis has been supported by inter-species comparative studies (Blumstein & Armitage, 1997; Bouchet, Blois-Heulin, & Lemasson, 2013; Krams, Krama, Freeberg, Kullberg, & Lucas, 2012; Manser et al., 2014; K. McComb & Semple, 2005; Wilkinson, 2003). However, while correlations across species are an important starting point for understanding vocal evolution, they do not tell us how complex vocalizations help animals navigate their social environment. Without this critical piece, our understanding of the adaptive nature of vocal complexity remains very limited. In other words, what we need know is, what is the *specific* function of vocal complexity in species shown to have comparatively high levels of sociality?

The first step to determining the specific functions of vocal complexity in highly social species is to identify the types of social challenges they need to be overcome. It is thought that species living in large social groups face high levels of conspecific noise and frequent separation with social partners in dense aggregations of conspecifics (Bee & Micheyl, 2008; H Brumm & Zollinger, 2011; Henrik Brumm & Slabbekoorn, 2005; Richardson & Lengagne, 2010; Schmidt & Römer, 2011). Complex vocalizations may function to counteract masking from conspecific noise or to help reach social partners outside of acoustic and/or visual contact of the caller. Altering vocalizations as a counterstrategy to noise occurs in anurans, marmosets, and humans, and several species produce salient ‘contact calls’ thought to help maintain intra-group spacing (Bee, 2007; Bee & Micheyl, 2008; Gustison & Townsend, 2015; Reichard & Welklin, 2014; Roy, Miller, Gottsch, & Wang, 2011). For species living in highly gregarious social groups, one of the challenges individuals face is in forming and maintaining social bonds (Curley & Keverne, 2005; Dunbar & Shultz, 2010). Complex vocalizations may serve to facilitate social bonding by promoting affiliative behaviors, such as close proximity, mating and grooming. Male songbirds, for example, are well-known to produce highly diverse sequences of notes that attract females and facilitate mating and the formation of pair bonds (B. E. Byers & Kroodsma, 2009; Kagawa & Soma, 2013; Vallet, Beme, & Kreutzer, 1998). Papionin primates frequently produce ‘grunt’ vocalizations that facilitate approaches and grooming interactions between social partners (Cheney & Seyfarth, 1997; Cheney, Seyfarth, & Silk, 1995; Palombit, Cheney, & Seyfarth, 1999; Rendall, Seyfarth, Cheney, & Owren, 1999; Whitham, Gerald, & Maestripieri, 2007).

An excellent species in which to study the specific social functions of complex vocalizations are geladas (*Theropithecus gelada*) – a non-human primate known for its high levels of sociality and unique vocal abilities compared to closely related species like baboons (Aich, Moos-Heilen, & Zimmermann, 1990; Bergman, 2013; Gustison, le Roux, & Bergman, 2012; Richman, 1976, 1987). Geladas live in large multi-level communities that are composed of smaller aggregations of harem-like reproductive units (Snyder-Mackler, Beehner, & Bergman, 2012). Two features of gelada sociality set them apart from other Papionins. First, within reproductive units, leader males form long-term (up to 8 years), affiliative bonds with the females in their unit (Gustison et al., 2012; Snyder-Mackler, Alberts, & Bergman, 2012). Second, reproductive units are found in dense aggregations ('herds') with 1-100 other units (occasionally >1,000 individuals, typically 200-500) (Snyder-Mackler, Beehner, et al., 2012). Thus, geladas are known for their high levels of sociality both in-terms of high levels of cross-sex bonding and the enormous numbers of individuals they encounter daily.

Geladas have also been publicized for their vocal complexity since the first studies in the 1970s (Dunbar & Dunbar, 1975; Richman, 1976). Recently, comparative data have shown that geladas have an expanded vocal repertoire because they produce a suite of low-amplitude call types not found in their baboon relatives (Gustison et al., 2012). Most of these phylogenetically "derived" calls – exhaled moans, wobbles, and vocalized yawns – are produced almost exclusively by adult males and contain highly salient acoustic properties (Gustison & Bergman, In Review, 2016). Exhaled moans are long in duration, wobbles contain rhythmic fluctuations in fundamental and formant frequencies, and yawns cover a wide frequency bandwidth. Gelada males almost always produce these derived call types in combination with exhaled grunts (a call type shared with baboons and many other primates) to form elaborate sequences that vary in composition and length (Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016). In terms of the function, complex sequences are generally used in close-range interactions between males and females (Gustison & Bergman, 2016; Gustison et al., 2012). However, why geladas need strings of multiple call types in situations that baboons use only grunts (Cheney et al., 1995; Palombit et al., 1999) remains a puzzle.

Here, we test the non-exclusive hypotheses that gelada vocal sequence complexity has a specific function that corresponds to (a) large group size or (b) long-term male-female social bonds. To test these hypotheses, we first identify which factors – social environment, social behaviors, and recipient type – determine the production and complexity of sequences. Then, we examine recipient responses to these sequences. Given their large group size, geladas frequently experience a social environment that is noisy and crowded (Gustison et al., 2012). Thus, longer vocal sequences or those containing salient derived call types may function to help males contact others in their reproductive units during conditions with limited acoustic and/or visual contact. If so, male vocal sequences should be produced at a higher rate, be longer in length, and/or contain more derived call types when conspecific noise level is high, unit females are farther away, and/or the level of conspecific density is high. It would also be expected that sequences are used during behavioral states in which individuals are out of contact with each other – like foraging and traveling.

On the other hand, geladas are characterized by stable long-term male-female social bonds that develop during a male's reproductive tenure (Gustison et al., 2012). Thus, complex vocal sequences may function to facilitate the social bonds that males maintain with unit females. As initial support for this hypothesis, a playback study showed unit females attend more strongly to male vocal sequences that contain a derived call (Gustison & Bergman, 2016). While suggestive that complex sequences are at least more salient to females, we still do not know how they function in natural male-female interactions. If longer complex sequences promote bond maintenance, male sequences that are longer in length or contain derived call types should be preferentially used during affiliative social behaviors like approaches and grooming. We also expect that males would produce these sequences specifically around unit females. Finally, we examine the responses of recipients to these sequences. If these complex vocal sequences function to facilitate social bonding, we would expect that they promote affiliative responses from females like approaches and grooming.

Methods

Study site and subjects

Data for this study come from adult male geladas across three different bands in one wild community (about 1,200 individuals) living in the Sankaber area of the Simien Mountains National Park, Ethiopia (Snyder-Mackler, Beehner, et al., 2012). Gelada bands are comprised of smaller harem-like units with a leader male, 0-3 subordinate follower males, and 1-11 females with their immature offspring (Snyder-Mackler, Beehner, et al., 2012). This population has been under intensive behavioral study since January 2006, and study subjects are fully habituated to human observers on foot (approach distances less than 3 m) (Bergman, 2010). The gelada research was approved by the University Committee on Use and Care of Animals at the University of Michigan and adhered to the laws and guidelines of Ethiopia.

Behavioral observations

From January to June 2014, we collected 251 15 minute focal samples from 13 leader and 6 follower males (1-20 focals per male) (Altmann, 1974). These samples included continuous monitoring of behavioral states (foraging, traveling, resting, socializing) and all-occurrence recording of affiliative behaviors. These social events included approaches, ano-genital inspection, copulation, grooming, infant handling, lip-smacking, and vocalizations. Vocalizations were also recorded with acoustic recording equipment (see below). Scan samples were taken at 5 minute intervals throughout the focal samples (n = 1,002 samples, 3-4 per focal). During scans, we recorded the focals' distance to the nearest same-unit adult female, as well as the number of other-unit adult males within a 10 m radius to the focal. The latter measurement was used as an index of conspecific density. The number of non-unit males within 10m to focal males during scans averaged 4.20 ± 0.15 (mean ± SEM) males (range: 0 to 26 males). The nearest unit female distance during scans averaged 2.75 ± 0.15m (range: 0 to 60m).

To determine precise timing latencies between vocal sequences and behavioral states, we collected 400 3-5 min video samples of 16 leader and 4 follower males using Flip UltraHD and Kodak Playsport video camcorders. These videos were taken during February 2013 to June 2014. Videos were taken *ad libitum* during a range of behavioral activities. These videos were labeled for behaviors using the same focal sampling protocol, and the start times of focal males' vocal sequences were recorded. We examined latencies between sequences and behavioral states that occurred within 5 s of each other. These videos also were used to determine the inter-call interval threshold that defined the boundary of a 'vocal sequence' (see Supplementary Materials).

Acoustic recordings

All occurrences of exhaled grunts, inhaled grunts, exhaled moans, inhaled moans, wobbles and vocal yawns were recorded during focal sampling (n = 6,195 individual calls from 1,406 sequences, 8-150 sequences per male) using a Sennheiser ME66 directional microphone connected to a Marantz PMD 661 Digital Recorder. These call types are combined together to make sequences, and have been described in previous studies (Gustison & Bergman, 2016; Gustison et al., 2012, 2016). We define a vocal sequence as one or more discrete calls made by the same individual and, for those sequences with two or more calls, separated by an inter-call interval of less than 2.983 s. This cut-off threshold was determined by creating a density plot of log-transformed inter-call interval durations of the six call types produced during 400 continuous 3-5 minute video recordings (see Supplemental Materials). Avisoft SAS Lab Pro (version 5.1.12, Avisoft Bioacoustics) was used to make spectrograms with a Fast Fourier transformation (size of 1,024 points) and label the calls in each sequence. The researcher labeling these recordings was blind to the male identity and behavior during the corresponding focal sample.

To determine conspecific noise level, 10 s acoustic recordings were taken at each 5 min scan sample. These recordings were made at the same volume level, with the microphone pointed in the direction of the focal male. A researcher blind to the focal male and behavioral events surrounding the scans labeled the start and end times of all gelada vocalizations in the scans using Praat (version 5.2.29). Spectrograms in Praat were viewed with a 0-6000 Hz range, 0.025 s window length, and 50 dB dynamic range. Vocalizations had to be visible on the spectrograms to be counted. Conspecific noise level was defined as the percentage of each scan that contained gelada calls. Conspecific noise in the 10-second scan recordings averaged 34.04 ± 0.01% (range: 0% to 100%). Preliminary analysis using a Linear Mixed Model (LMM) showed that conspecific noise was related to other social environment variables and focal behavioral state during the scan samples. In this LMM, conspecific density, female distance, and behavioral state were fixed factors, and the focal sample and male identity were random effects. We found that the percent of conspecific noise was negatively associated with unit female distance (Intercept = 28.783 ± 2.585, Estimate (± SEM) = -0.492 ± 0.183, $t = -2.69$, $p = 0.0074$) and positively associated with the number of non-unit males within 10 m (Estimate = 2.142 ± 0.220, $t = 9.752$, $p < 0.0001$). Conspecific noise during resting (45.54%) and socializing (42.27%) was higher than during foraging (26.38%; rest vs. forage: $t = 2.27$, $p = 0.0236$; social vs. forage: $t = 2.12$, $p = 0.0339$) and traveling (26.44%; rest vs. travel: $t = 2.21$, $p = 0.0272$; social vs. travel: $t = 2.05$, $p = 0.0408$). Conspecific noise during resting and socializing did not differ ($p > 0.05$), nor did it differ during foraging and traveling ($p > 0.05$).

Data analysis

Six dependent variables were used to represent sequence production and complexity. ‘Production’ was quantified by the sequence rate and sequence size (number of calls per sequence). The time window used to calculate sequence rate depended on the analysis. Sequence ‘complexity’ was quantified by the repertoire size of derived calls (exhaled moans, wobbles, and vocalized yawns) and the total number of each type of derived call per vocal sequence. A series of LMMs were used to test the causes and consequences of sequence production and complexity. All LMMs included focal sample and male identity as random effects.

First, we tested the effects of social environment on sequence production and complexity. Social environment fixed factors included conspecific noise, conspecific density, and female distance. These factors were measured for each 5 min focal sample interval (n = 753 intervals) by averaging the values from the two adjacent scan samples. Then, we ran six LMMs, each with the averages of conspecific noise, conspecific density, and female distance as fixed factors. The behavioral states of the first and second adjacent scan samples were included as random effects. The first LMM tested whether the sequence rate for each 5 min interval was predicted by any of the three social environment variables. The next five LMMs tested whether sequence size or complexity was predicted by the social environment variables for the 5 min intervals that the sequences (n = 1,406) were produced in.

Second, we tested the effects of male behavioral states on sequence production and complexity. We separated each behavioral state into two time intervals, a transition period (5s before to 5s after the start of the current state) and a non-transition period (5s after the start of the current state to 5s before the start of the next state). Then, we ran six LMMs, each with behavioral state and time interval as the fixed factor. The first LMM tested whether the sequence rate (per 10 s) was predicted by the corresponding state and its time interval. For this model, only behavior states that (a) lasted at least 20 s and (b) had at least 5s of observation before they started were used so that there was a minimum of 10s per time interval for each state. Thus, we examined sequence rates in the transition and non-transition periods. The next five LMMs tested whether sequence size or complexity was predicted by the behavioral state/time interval that each sequence occurred in.

Third, we tested the effects of specific social behaviors on sequence production and complexity. We first used LMMs to test whether sequence rate was higher in the 10 s window during the start of each social behavior compared to the 15-5 s before and 5-15 s after the behavior started. These LMMs included time interval (before, during, after) as a fixed factor. Then, we used LMMs to test whether the size or complexity of 'non-social' sequences produced during the 10 s window of a social behavior differed from those produced outside the 10 s window. In this analysis, comparisons were only made for social behaviors for which sequence rate was enhanced.

Fourth, we tested the effect of receiver type on sequence production and complexity. It is often difficult to tell who the receivers are during a vocalization, and so we focused on sequences given within the 10 s window surrounding approach events. Specifically, we compared approaches that either were or were not towards an adult unit female (the fixed factor in LMMs). We used an LMM to test whether sequence rate differed between approaches with a unit female or other individuals. Then, we used LMMs to test whether the size or complexity of sequences produced during approach with unit females differed from approaches with other individuals.

Last, we tested the potential influence of sequence production and complexity on receiver responses. We examined two types of responses – (a) change in unit female proximity to the focal males and (b) elicitation of grooming behavior from a unit female following an approach. The first LMM included sequence rate (per 5 min focal interval) as a dependent variable and the changes in nearest unit female and non-unit male distances (scan sample 2 minus scan sample 1 distances) as fixed factors. Non-unit male distances were included in the model to control for potential changes in overall herd density. The next five LMMs included sequence size or complexity measures as the dependent variable and changes in unit-female and non-unit male distances in the corresponding focal interval as fixed effects. Then, we ran an LMM that included sequence rate (per 10 s) during an approach with a unit female as the dependent variable and grooming outcome that occurred within 10 s of the approach as the fixed factor. The next five LMMs included size or complexity measures as the dependent variable and grooming outcome as the fixed factor.

LMMs were run using the function “lmer” of the R package “lme4” (Bates, Maechler, & Bolker, 2012). The “lmerTest” package was implemented to determine the significance of the LMM coefficients (Kuznetsova, Brockhoff, & Christensen, 2013). We used the “difflsmeans” function in the lmerTest package to make pairwise comparisons. All calculations and statistical tests were carried out in R 3.3.0 (R Development Core Team, 2016). Critical values were set at alpha = 0.05, and all tests were two-tailed.

Results

Gelada male vocal sequences were produced at a mean (± SEM) rate of 1.87 ± 0.06 (range: 0-10) sequences per 5 min and had a mean size of 4.41 ± 0.10 (1-23) calls. Sequences had a mean repertoire of 0.36 ± 0.02 (0-3) derived call types and contained 0.27 ± 0.01 (0-3) exhaled moans, 0.05 ± 0.01 (0-6) wobbles, and 0.08 ± 0.01 (0-1) yawns. First, we explored the effect of social environment, male behavior and receiver type on sequence production and complexity. Then, we explored receiver responses to the production and complexity of sequences.

Effect of social environment

We used LMMs to test the potential effects of conspecific noise, conspecific density, and unit female distance on vocal sequence production and complexity (Table 6.1). Conspecific noise level was positively associated with sequence rate. Conspecific noise level also was positively associated with the number of derived call types per sequence (Fig 6.1a), and the number of exhaled moans per sequence (Fig 6.1b). There was no evidence that conspecific noise level was associated to sequence size or the number of wobbles and yawns per sequence. Conspecific density was negatively associated with sequence rate and positively associated with the number of yawns per sequence (Fig 6.1c). There was no evidence that conspecific density was associated to the number of derived call types per sequence or the number of exhaled moans and wobbles per sequence. There also was no evidence that nearest unit female distance was associated to the any of the dependent variables.

Effect of male behavior

First, we used LMMs to test the potential influence of behavioral states (foraging, n = 562, traveling, n = 367, resting, n = 246, and socializing, n = 269) and the corresponding time interval (transition or non-transition periods) on sequence production and complexity. Pairwise comparisons showed that sequence rate was highest during transitions to socializing compared to all other state/time interval categories ($p < 0.0001$ for all comparisons, Fig 6.2a). The only other difference found was that sequence rate was higher during foraging transitions than during non-transition foraging periods ($t = 2.44$, $p = 0.0148$; Fig 6.1a). Then, we examined the latency between sequence production in the transition periods of foraging and socializing states from the video samples. On average, sequences occurred about 0.59 ± 0.38 s after a foraging state ($n = 61$ latencies) started and 1.12 ± 0.35 s before the start of a socializing state ($n = 52$ latencies; Fig 6.2b). There were subtle differences in sequence size and complexity measures across behavioral state/time intervals (Table 6.2). Sequence size was smallest in non-transition foraging and non-transition socializing periods. Derived call repertoire size per sequence was highest during resting and socializing. The number of exhaled moans per sequence was highest during resting and transitions to socializing. The number of wobbles per sequence was highest during transitions to socializing. The number of yawns per sequence was highest during resting and transitions to socializing. We also explored the sequence structure in the four behavioral states (see Supplementary Material). Sequence structure was not remarkably different between the states and all sequence types had transitional probabilities that differed from what was expected due to chance. Most sequences began with exhaled grunts or exhaled moans and ended with inhaled moans or vocalized yawns. Sequences also contained frequent alterations between exhaled and inhaled grunts.

Next, we compared sequence production and complexity across different types of affiliative social behaviors (give and receive approaches, give ano-genital inspections, give copulations, give and receive grooming, infant handling, and give and receive lip-smacking). First, we carried out a suite of LMMs to test whether sequence rate was higher during the 10 s window during the start of a social behavior compared to the 15-5 s before and 5-15 s after (Table 6.3, Fig 6.3). Sequence rates were higher during approaches (give = 350 approaches, receive = 374 approaches; Fig 6.3a-b) and all other social behaviors except for when males received lip-smacking. Also, sequence size and complexity differed between non-social sequences and those produced during social behaviors (Table 6.4, Fig 6.3). Sequence size was larger in sequences produced during all social behaviors except for when males received grooming. The derived call repertoire size of sequences was larger when males gave and received approaches, engaged in ano-genital inspection, received grooming, and gave a lip-smack. The number of exhaled moans per sequence was only greater in sequences produced when males gave an approach. The number of wobbles per sequence was greater during all social behaviors expect for infant handling. The number of yawns per sequence was greater when males gave an approach, received grooming, and handled infants. The number of yawns per sequence was lower when males gave grooming.

Effect of receiver type

We examined how sequence production and complexity differed based on whether the recipient was an adult female from the focal male's unit. First, we tested whether sequence rate was greater during an approach to a unit female (n = 490 approaches) compared to others (n = 252). We found that sequence rate was higher during an approach to a unit female than another individual (t = 2.77, p = 0.0057; Fig 6.4a). Next, we tested whether sequences produced during an approach differed in size or complexity when the approach was with a unit female (n = 192 sequences) compared to others (n = 71). We found that sequences were larger in size if they were directed at a unit female (t = 2.26, p = 0.0249; Fig 6.4b), but they did not differ in their derived call repertoire size (t = 0.52, p = 0.6027; Fig 6.4c). There was also no evidence that sequences differed in the number of exhaled moans (unit female: 0.40 ± 0.04 moans, other: 0.21 ± 0.03 moans; t = 1.52, p = 0.1290), wobbles (unit female: 0.17 ± 0.05 wobbles, other: 0.14 ± 0.04 wobbles; t = -0.40, p = 0.6908), or yawns (unit female: 0.14 ± 0.03 yawns, other: 0.18 ± 0.03 yawns; t = -0.84, p = 0.4024).

Receiver responses to sequences

We examined whether sequence production and complexity played a potential role in influencing receiver proximity to the focal and affiliative behavior. First, we examined changes in unit female and non-unit male distances to the focal across 5 min intervals (753 intervals, 1,406 sequences). As expected, we found that unit female distances decreased as sequence rate increased (Intercept (± SEM) = 1.867 ± 0.131, Estimate (± SEM) = -0.041 ± 0.014, t = -2.97, p = 0.0031), and there was no clear change in non-unit male distances based on sequence rate (Estimate = -0.004 ± 0.011, t = -0.32, p = 0.7517). Sequence size did not predict changes in unit female distance (Intercept = 4.140 ± 0.310, Estimate = 0.011 ± 0.024, t = 0.45, p = 0.6497) or non-unit male distance (Estimate = -0.027 ± 0.018, t = -1.5, p = 0.1341). Derived call repertoire size per sequence did not predict changes in unit female distance (Intercept = 0.365 ± 0.046, Estimate = -0.001 ± 0.004, t = -0.15, p = 0.877) or non-unit male distance (Estimate = 0.000 ± 0.003, t = 0.12, p = 0.9079). The number of exhaled moans per sequence did not predict changes in unit female distance (Intercept = 0.272 ± 0.042, Estimate = -0.002 ± 0.003, t = -0.7, p = 0.4815) or non-unit male distance (Estimate = 0.004 ± 0.002, t = 1.74, p = 0.0822). The number of wobbles per sequence did not predict changes in unit female distance (Intercept = 0.049 ± 0.016, Estimate = 0.002 ± 0.002, t = 0.91, p = 0.3656) or non-unit male distance (Estimate = 0.000 ± 0.002, t = -0.28, p = 0.7802). The number of yawns per sequence did not predict changes in unit female distance (Intercept = 0.092 ± 0.012, Estimate = -0.001 ± 0.002, t = -0.37, p = 0.7121) or non-unit male distance (Estimate = -0.002 ± 0.001, t = -1.28, p = 0.1999).

Then, we examined whether sequence production or complexity during an approach with a unit female was predicted by the grooming outcome (no grooming, males giving grooming, males receiving grooming). First, we compared sequence rates during approaches followed by no grooming (n = 372 approaches) to approaches followed by males giving grooming (n = 52) and males receiving grooming (n = 49). We found that sequence rate was higher if males gave (t = 6.99, p < 0.0001) or received grooming (t = 6.44, p < 0.0001) compared to if there was no grooming (Fig 6.5a). Then we compared the size and complexity of approach sequences followed by no grooming (n = 109 sequences), males giving grooming (n = 43) and males receiving grooming (n = 37). We found no evidence that sequence size differed if males gave (t = 0.24, p = 0.8126) or received (t = -0.31, p = 0.7565) grooming compared to no grooming (Fig 6.5b). We also found no evidence that the derived call repertoire differed if males gave grooming (t = -0.62, p = 0.5385) compared to no grooming, however, derived call repertoire size was larger if males received grooming (t = 2.05, p = 0.0421) compared to no grooming (Fig 6.5c). We found no evidence that the number of exhaled moans per sequence differed if males gave grooming (t = -0.73, p = 0.4655) or received grooming (t = -0.22, p = 0.8292) compared to no grooming (Fig 6.5d). The number of wobbles per sequence was greater if males gave grooming (t = 2.64, p = 0.0091) compared to no grooming, but there was no evidence that wobble number differed if males received grooming (t = 0.27, p = 0.7858) compared to no grooming (Fig 6.5e). Finally, sequences were less likely to contain a yawn if males gave grooming (t = -2.28, p = 0.0239) and more likely to contain a yawn if males received grooming (t = 3.45, p = 0.0007) compared to no grooming (Fig 6.5f).

Discussion

In this study, we show that complex gelada vocal sequences have specific functions that tie into two social complexity metrics – large group size and high levels of socializing. First, we found evidence that complex vocal sequences function to combat the conspecific noise created by living in a large and crowded social group. Second, we found evidence that complex vocal sequences facilitate prosocial interactions between adult males and females.

One of the challenges vertebrates face when in proximity to several conspecifics is noise level, and counterstrategies to noise include diversifying vocalizations to be more salient as well as making vocalizations redundant through repetition or prolongation (Bee & Micheyl, 2008; Henrik Brumm & Slabbekoorn, 2005; Hotchkin & Parks, 2013; Richardson & Lengagne, 2010; Schmidt & Römer, 2011). We found evidence for both counterstrategies. Gelada males diversified their vocal sequences by using a larger repertoire of derived call types during periods of high noise levels. This finding mirrors analogous work on the choruses of male Túngara frogs (*Physalaemus pustulosus*), which shows that males produce higher proportions of complex sequences (i.e., those that include 'chuck' vocalizations) when chorusing in higher male densities, and therefore higher noise, conditions (Bernal, Page, Rand, & Ryan, 2007). This finding also supports the interpretations of a previous playback experiment showing that female geladas respond more strongly to vocal sequences containing derived call types – it was proposed that this result occurred because of the higher salience of derived calls (Gustison & Bergman, 2016). Second, gelada males engaged in more redundant vocal behaviors during high noise levels. They produced vocal sequences at higher rates and produced more exhaled moans, a derived call type known for its extended duration (Gustison & Bergman, In Review, 2016). Such findings support studies of captive marmosets (*Callithrix jacchus*), which are thought to produce prolonged phee calls during noise to enhance signal transmission (Brumm, 2004).

Overall, we found little evidence to suggest that gelada male vocal sequences function to counteract other challenges of living in a large group, like limited visual contact with unit females or needing to contact several individuals simultaneously. Sequence rate and derived call repertoire size were unrelated to conspecific density, and sequences were produced at low rates during foraging and traveling, which is when individuals would be most likely to be out of visual contact with social partners. The one exception to this rule was the production of vocal yawns. We found that vocal sequences were more likely to contain vocal yawns when male density was high, a finding that supports previous captive gelada work suggesting that vocal yawns may serve as a multi-model agonistic displays between males (Leone, Ferrari, & Palagi, 2014). Taken together, these contrasting associations between specific derived call types and social environment (exhaled moans with conspecific noise and vocal yawns with conspecific density) suggest that different selective pressures associated with large group size may have promoted the development of different derived call types in geladas.

Apart from the several avian taxa that express monogamy, highly gregarious and long-term male-female social bonds are relatively rare among vertebrates (Greenwood, 1980; Kleiman, 1977). One of the challenges faced by males in such a position is how to effectively form and maintain their social bonds with reproductive females. Male vertebrates commonly use elaborate vocal displays to attract short-term female mating partners, which is characteristic of red deer (*Cervus elaphus*), túngara frogs (*P. pustulosus*) and canaries (*Serinus canaria*), for example (Charlton, Reby, & McComb, 2007; Drăgănoiu, Nagle, & Kreutzer, 2002; Kime, Rand, Kapfer, & Ryan, 1998; McComb, 1991; Vallet et al., 1998). Less understood is how males might use complex vocal strategies to develop long-term associations with females. Males of monogamous species – like zebra finches (*Taeniopygia guttata*) and California mice (*Peromyscus californicus*) – continue to direct vocalizations towards their female partners after a pair bond has formed (Pultorak, Fuxjager, Kalcounis-Rueppell, & Marler, 2015; Tomaszycki & Adkins-Regan, 2006). In humans, higher conversation quality and quantity with an intimate partner is associated with higher relationship satisfaction (E. S. Byers, 2005; Emmers-Sommer, 2004). The findings of the current study expand on these earlier insights into the role of vocal communication in maintaining cross-sex social bonds by showing that gelada males rely on vocal sequences, specifically those containing more derived call types, to facilitate their affiliative interactions with females. Importantly, the composition of gelada vocal sequences predicted the directionality of heterosexual social interactions. The presence of a sequence during an approach and its size (number of calls) equally predicted giving or receiving of grooming. However, diversified vocal sequences (i.e., more derived call types) and those containing vocal yawns predicted female to male grooming, while sequences containing more wobbles predicted male to female grooming. These vocal sequence types may therefore serve as referential signals in social contexts. This interpretation is supported by earlier experimental work in which female geladas chose to spend more time close to playbacks of male derived call vocal sequences than grunt-only sequences (Gustison & Bergman, 2016).

Previous comparative research in primates demonstrated that vocal repertoire size increased in taxa expressing larger group sizes and higher grooming rates (K. McComb & Semple, 2005). These conclusions left an open question about how increased levels of sociality promote the need for more communicative complexity. Comparing the current study with previous baboon work starts to answer this question by emphasizing the significance of conspecific noise and density levels of large groups and the need to effectively maintain long-term heterosexual social bonds. A previous study showed that geladas range in larger aggregations and vocalize at higher rates than chacma baboons (*Papio ursinus*), suggesting that conspecific noise is a factor promoting gelada vocal complexity (Gustison et al., 2012). The current findings support this interpretation by showing that sequences containing more derived call types – specifically exhaled moans – occur in noisier social environments, and sequences containing vocal yawns are more prevalent in high density environments. On the other hand, geladas and baboons have contrasting social systems that can account for differences in communicative abilities. Most baboons live in multi-male multi-female polygamous social groups in which males form temporary consortships and ‘friendships’ with females (Clutton-Brock, T. H., Harvey, 1977; Palombit et al., 1999). Similar to geladas, baboon males who approach female ‘friends’ while grunting were more likely to engage in grooming and other affiliative interactions than if approaches did not co-occur with grunting (Palombit et al., 1999). A key difference, however, is that gelada males maintain intimate relationships with multiple females, and appear to use derived vocal sequences as rudimentary referential signals that predict the directionality of affiliative interactions. By having a more diverse way of communicating, gelada males appear better able to elicit female investment in the relationship. It is also possible that gelada males can reach several unit members simultaneously when they use vocal sequences, thereby maintaining multiple social bonds in a single vocal utterance. This strategy – termed ‘vocal grooming’ – is a well-known explanation for the high vocal rates in humans and geladas (Dunbar, 1998, 2003).

Taken together, the findings show that the need to maintain bonds within a noisy backdrop of conspecific vocalizations may favor greater vocal complexity. Future work will be needed to examine why it is that gelada vocal complexity is significant for cross-sex bonds but not in same-sex female bonds, even though both types occur in a noisy environment. One possibility is that males need to form new social bonds, whereas females are born into their social bonds, suggesting that vocal complexity may help accelerate a male’s integration into a unit.

Acknowledgements

We thank the Ethiopian Wildlife Conservation Authority (EWCA) as well as the wardens and staff of the Simien Mountain National Park for permission and support in conduction research on geladas. We also thank all the members of the University of Michigan Gelada Research Project (UMGRP) for assistance in data collection and valuable insight on analyses and interpretation. MLG was supported by grants from the National Geographic Society (9122-12, W304-14), Leakey Foundation, University of Michigan, and Petridish (primary supporter: John Allen). UMGRP was supported by the Wildlife Conservation Society (SSF 67250), the National Geographic Society (8100-06), the Leakey Foundation, the National Science Foundation (BCS-0715179, BCS-0962118, IOS- 1255974), and the University of Michigan.

Supplementary Materials

Determination of vocal sequence boundaries

From February 2013 to June 2014, 400 3-5 min video recordings were taken on 20 males using Flip UltraHD video camcorder and Kodak Playsport video camcorders. We used Audacity (version 2.0.5) to extract the audio from these recordings and Avisoft SAS Lab-Pro (version 5.1.12, Avisoft Bioacoustics) to label the start and end times of exhaled grunts, inhaled grunts, exhaled moans, inhaled moans, wobbles and yawns (5,218 total calls) produced by the focal males. Then, we measured the inter-call intervals (n = 4,868 intervals) between calls produced during the same videos. Following previously published methods (Kershenbaum, Ilany, Blaustein, & Geffen, 2012), we chose the cut-off threshold for interval length based on the lowest point between the first two peaks of a density plot of log-transformed inter-call interval lengths (Fig 6.S1). This analysis showed that 2.983 s was an appropriate threshold to split adjacent calls into different sequences.

Sequence structure across behavioral states

We performed lag sequential analysis (Bakeman & Gottman, 1986; Bakeman & Quera, 1992) to characterize patterns of call transitions in gelada male vocal sequences using the GSEQ program 5.1 (Bakeman & Quera, 1992). Lag sequential analysis shows whether the association between any two calls is non-random by testing whether the transitions occur at rates significantly different than expected given the observed frequency of calls. We examined the frequency of calls at Lag1 (the call immediately following the initial call, Lag0). All calls, including ‘silence’ (indicating the start or end of a sequence), were used as initial and following calls, producing a 7x7 table of p-values for the conditional probabilities. Calls were not allowed to repeat (i.e., a repeated call was counted once) so that the diagonal of the transition matrix did not contain any structural zeros (Bakeman & Gottman, 1986). We used a chi-squared test to examine whether the observed frequency table differed from the expected frequency table. When analyzing specific call transitions, we used an alpha value of 0.05.

We looked at sequences produced during the different behavioral states, foraging (2,554 calls from 627 sequences across 18 males), traveling (988 calls from 181 sequences across 17 males), resting (831 calls from 174 sequences across 16 males) and socializing (1,822 calls from 424 sequences across 19 males) separately. Lag sequential analysis suggested that call transitions in gelada sequences are non-independent. We found that transition frequencies deviated from random expectations for foraging sequences (χ^2 = 1,047.496, df = 29, $p < 0.0001$), traveling sequences (χ^2 = 439.114, df = 29, $p < 0.0001$), resting sequences (χ^2 = 277.218, df = 29, $p < 0.0001$), and socializing sequences (χ^2 = 501.468, df = 29, $p < 0.0001$). Further analysis of these transitions revealed that certain call types tended to start or end sequences, groups of calls tended to occur together, and patterns were similar across states (Fig 6.S2). Sequences were typically begun with exhaled grunts or exhaled moans and ended by inhaled moans and yawns. Exhaled and inhaled grunts were often produced in alternating cycles.

References

Aich, H., Moos-Heilen, R., & Zimmermann, E. (1990). Vocalizations of adult gelada baboons (*Theropithecus gelada*): Acoustic structure and behavioural context. *Folia Primatologica*, *55*, 109–132.

Altmann, J. (1974). Observational study of behavior : Sampling methods. *Behaviour*, *49*, 227–267.

Baker, C. M. (1988). Vocalizations of Captive Water Mongooses, *Atilax-Paludinosus*. *Zeitschrift Fur Saugetierkunde*, *53*, 83–91.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. Retrieved from http://cran.r-project.org/ package=lme4

Bee, M. a. (2007). Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. *Animal Behaviour*, *74*, 549–558.

Bee, M. a, & Micheyl, C. (2008). The cocktail party problem: what is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology*, *122*, 235–51.

Bergman, T. J. (2010). Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 3045–53.

Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Current Biology*, *23*, R268-9.

Bernal, X. E., Page, R. A., Rand, A. S., & Ryan, M. J. (2007). Natural history miscellany - Cues for eavesdroppers: Do frog calls indicate prey density and quality? *The American Naturalist*, *169*, 409–415.

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

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: a comparison of three non-human primate species. *Frontiers in Psychology*, *4*, 390.

Brumm, H. (2004). Acoustic communication in noise: regulation of call characteristics in a New World monkey. *Journal of Experimental Biology*, *207*, 443–448.

Brumm, H., & Slabbekoorn, H. (2005). Acoustic Communication in Noise. *Advances in the Study of Behavior*, *35*, 151–209.

Brumm, H., & Zollinger, S. A. (2011). The evolution of the Lombard effect: 100 years of psychoacoustic research. *Behaviour*, *148*, 1173–1198.

Byers, B. E., & Kroodsma, D. E. (2009). Female mate choice and songbird song repertoires. *Animal Behaviour*, *77*, 13–22.

Byers, E. S. (2005). Relationship satisfaction and sexual satisfaction: a longitudinal study of individuals in long-term relationships. *Journal of Sex Research*, *42*, 113–118.

Charlton, B. D., Reby, D., & McComb, K. E. (2007). Female red deer prefer the roars of larger males. *Biology Letters*, *3*, 382–385.

Cheney, D. L., & Seyfarth, R. M. (1997). Reconciliatory grunts by dominant female baboons influence victims’ behaviour. *Animal Behaviour*, *54*, 409–18.

Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Animal Behaviour*, *50*, 249–257.

Clutton-Brock, T. H., Harvey, P. H. (1977). Primate ecology and social organization. *Journal of Zoology*, *183*, 1–39.

Curley, J. P., & Keverne, E. B. (2005). Genes, brains and mammalian social bonds. *Trends in Ecology and Evolution*, *20*, 561–567.

Drăgănoiu, T. I., Nagle, L., & Kreutzer, M. (2002). Directional female preference for an exaggerated male trait in canary (*Serinus canaria*) song. *Proceedings of the Royal Society B: Biological Sciences*, *269*, 2525–31.

Dunbar, R. I. M. (1998). Grooming, Gossip, and the Evolution of Language. In *Grooming, Gossip, and the Evolution of Language*. Cambridge, MA: Harvard University Press.

Dunbar, R. I. M. (2003). The social brain: Mind, language, and society in evolutionary perspective. *Annual Reviews in Anthropology*, *32,* 163-181.

Dunbar, R. I. M., & Dunbar, P. (1975). *Social Dynamics of Gelada Baboons*. Basel, Switzerland: Karger.

Dunbar, R. I. M., & Shultz, S. (2010). Bondedness and sociality. *Behaviour*, *147*, 775–

803.

Emmers-Sommer, T. M. (2004). The effect of communication quality and quantity indicators on intimacy and relational satisfaction. *Journal of Social and Personal Relationships*, *21*, 399–411.

Freeberg, T. M. (2006). Social complexity can drive vocal complexity: group size influences vocal information in Carolina chickadees. *Psychological Science*, *17*, 557–61.

Greenwood, P. J. (1980). Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, *28*, 1140–1162.

Gustison, M. L., & Bergman, T. J. (In Review). Divergent acoustic properties of gelada and baboon vocalizations and their implications for the evolution of human speech. *Journal of Language Evolution*.

Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Scientific Reports*, *6*, 19680.

Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1847–59.

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*, E2750–E2758.

Gustison, M. L., & Townsend, S. W. (2015). A survey of the context and structure of high- and low-amplitude calls in mammals. *Animal Behaviour*, *105*, 281–288.

Hotchkin, C., & Parks, S. (2013). The Lombard effect and other noise-induced vocal modifications: insight from mammalian communication systems. *Biological Reviews*, *88*, 809–24.

Kagawa, H., & Soma, M. (2013). Song performance and elaboration as potential indicators of male quality in Java sparrows. *Behavioural Processes*, *99*, 138–144.

Kime, N. M., Rand, A. S., Kapfer, M., & Ryan, M. J. (1998). Consistency of female choice in the túngara frog: a permissive preference for complex characters. *Animal Behaviour*, *55*, 641–649.

Kleiman, D. G. (1977). Monogamy in mammals. *The Quarterly Review of Biology*, *52*, 39–69.

Krams, I., Krama, T., Freeberg, T. M., Kullberg, C., & Lucas, J. R. (2012). Linking social complexity and vocal complexity: a parid perspective. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1879–91.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/package=lmerTest

Leone, A., Ferrari, P. F., & Palagi, E. (2014). Different yawns, different functions? Testing social hypotheses on spontaneous yawning in *Theropithecus gelada*. *Scientific Reports*, *4*, 4010.

Manser, M. B., Jansen, D. A. W. A. M., Graw, B., Hollén, L. I., Bousquet, C. A. H., Furrer, R. D., & Le Roux, A. (2014). Vocal complexity in meerkats and other mongoose species. *Advances in the Study of Behavior*, *46*, 281–310.

McComb, K. E. (1991). Female choice for high roaring rates in red deer, *Cervus elaphus*. *Animal Behaviour*, *41*, 79–88.

McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, *1*, 381–5.

Palombit, R. A., Cheney, D. L., & Seyfarth, R. M. (1999). Male grunts as mediators of social interaction with females in wild chacma baboons (*Papio cynocephalus ursinus*). *Behaviour*, *136*, 221–242.

Pultorak, J. D., Fuxjager, M. J., Kalcounis-Rueppell, M. C., & Marler, C. A. (2015). Male fidelity expressed through rapid testosterone suppression of ultrasonic vocalizations to novel females in the monogamous California mouse. *Hormones and Behavior*, *70*, 47–56.

R Development Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Reichard, D. G., & Welklin, J. F. (2014). On the existence and potential functions of low-amplitude vocalizations in North American birds. *The Auk*, *132*, 156–166.

Rendall, D., Seyfarth, R. M., Cheney, D. L., & Owren, M. J. (1999). The meaning and function of grunt variants in baboons. *Animal Behaviour*, *57*, 583–592.

Richardson, C., & Lengagne, T. (2010). Multiple signals and male spacing affect female preference at cocktail parties in treefrogs. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 1247–52.

Richman, B. (1976). Some vocal distinctive features used by gelada monkeys. *The Journal of the Acoustical Society of America*, *60*, 718–24.

Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates*, *28*, 199–223.

Roy, S., Miller, C. T., Gottsch, D., & Wang, X. (2011). Vocal control by the common marmoset in the presence of interfering noise. *The Journal of Experimental Biology*, *214*, 3619–29.

Schmidt, A. K. D., & Römer, H. (2011). Solutions to the cocktail party problem in insects: selective filters, spatial release from masking and gain control in tropical crickets. *PloS One*, *6*, e28593.

Snyder-Mackler, N., Alberts, S. C., & Bergman, T. J. (2012). Concessions of an alpha male? Cooperative defence and shared reproduction in multi-male primate groups. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 3788–3795.

Snyder-Mackler, N., Beehner, J. C., & Bergman, T. J. (2012). Defining higher levels in the multilevel societies of geladas (*Theropithecus gelada*). *International Journal of Primatology*, *33*, 1054–1068.

Tomaszycki, M. L., & Adkins-Regan, E. (2006). Is male song quality important in maintaining pair bonds? *Behaviour*, *143*, 549–567.

Vallet, E., Beme, I., & Kreutzer, M. (1998). Two-note syllables in canary songs elicit high

Whitham, J. C., Gerald, M. S., & Maestripieri, D. (2007). Intended receivers and functional significance of grunt and girney vocalizations in free-ranging female rhesus macaques. *Ethology*, *113*, 862–874.

Wilkinson, G. S. (2003). Social and vocal complexity in bats. In F. B. M. de Waal & P. L. Tyack (Eds.), *Animal social complexity: Intelligence, culture, and individualized societies* (pp. 322–341). Cambridge, MA: Harvard University Press.

References – Supplementary Materials

Bakeman, R., & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge University Press.

Bakeman, R., & Quera, V. (1992). SDIS: A sequential data interchange standard. *Behavior Research Methods, Instruments, & Computers*, *24*, 554–559.

Kershenbaum, A., Ilany, A., Blaustein, L., & Geffen, E. (2012). Syntactic structure and geographical dialects in the songs of male rock hyraxes. Proceedings of the Royal Society, Series B: Biological Sciences, *279*, 2974–81.

Table 6.1. Results of Linear Mixed Models testing associations between vocal sequence production and complexity with features of the social environment.

Dependent variable	Intercept			Conspecific Noise (proportion of scan)			Conspecific Density (non-unit males in 10 m radius)			Unit Female Distance (m)		
	Estimate [SEM]	**t**	**p**	**Estimate [SEM]**	**t**	**p**	**Estimate [SEM]**	**t**	**p**	**Estimate [SEM]**	**t**	**p**
Sequence rate[1]	1.98 [0.304]	6.52	**0.0004**	0.811 [0.351]	2.31	**0.0210**	-0.061 [0.02]	-2.99	**0.0030**	0.000 [0.019]	0.01	0.9947
Sequence size[2]	4.092 [0.536]	7.63	**<0.0001**	0.664 [0.572]	1.16	0.2458	0.019 [0.032]	0.61	0.5409	-0.004 [0.038]	-0.11	0.9151
Derived call repertoire size[2]	0.295 [0.089]	3.33	**0.0127**	0.279 [0.089]	3.14	**0.0018**	0.005 [0.005]	1.10	0.2722	-0.006 [0.006]	-1.08	0.2804
Number of exhaled moans[2]	0.219 [0.051]	4.25	**0.0002**	0.273 [0.076]	3.60	**0.0004**	-0.003 [0.004]	-0.77	0.4443	-0.009 [0.005]	-1.88	0.0616
Number of wobbles[2]	0.015 [0.028]	0.52	0.6052	0.062 [0.05]	1.25	0.2116	0.004 [0.003]	1.64	0.1026	0.000 [0.003]	0.09	0.9263
Number of yawns[2]	0.081 [0.062]	1.30	0.2724	0.004 [0.041]	0.09	0.9262	0.009 [0.002]	3.76	**0.0002**	0.001 [0.003]	0.45	0.6538

Table 6.2. Comparisons between sequence size and complexity measures based on the corresponding behavioral state and time intervals. Descriptive statistics listed as mean [SEM], and values in the same row with no overlapping letters differ from each other ($p < 0.05$) based on pairwise comparisons using Linear Mixed Models.

Dependent variable	**Forage**		**Travel**		**Rest**		**Socialize**	
	Transition (n = 68)	**Non-transition (n = 559)**	**Transition (n = 82)**	**Non-transition (n = 99)**	**Transition (n = 29)**	**Non-transition (n = 145)**	**Transition (n = 135)**	**Non-transition (n = 289)**
Sequence size	5.07 [0.50] a,c	3.95 [0.14] b	4.90 [0.51] a,b	5.92 [0.45] c	5.48 [0.82] a,c	4.63 [0.33] a,d	5.39 [0.37] a,c	3.79 [0.17] a,d
Derived call repertoire size	0.25 [0.06] a	0.22 [0.02] a	0.23 [0.05] a	0.33 [0.06] a,b	0.72 [0.14] c,d	0.68 [0.07] c	0.50 [0.06] d,e	0.42 [0.04] b,e
Number of exhaled moans	0.22 [0.05] a	0.23 [0.02] a,c	0.20 [0.05] a,c	0.23 [0.05] a,c	0.45 [0.12] b	0.36 [0.05] b	0.31 [0.05] b,c,d	0.31 [0.03] a,b,d
Number of wobbles	0.00 [0.00] a,b,d	0.00 [0.00] a	0.01 [0.01] a,d	0.11 [0.05] b,d	0.00 [0.00] a,b,d	0.06 [0.03] a,b,d	0.21 [0.06] c	0.07 [0.02] d
Number of yawns	0.03 [0.02] a,d	0.01 [0.00] a	0.05 [0.02] a,d	0.04 [0.02] a,d	0.34 [0.09] b	0.33 [0.04] b	0.13 [0.03] c	0.09 [0.02] c,d

Table 6.3. Results of Linear Mixed Models comparing sequence rates before, during and after affiliative social behaviors.

		Sequence Rate per 10s (mean [SEM])			**During vs. Before LMM Results**		**During vs. After LMM Results**	
Social Behavior	**N**	**Before**	**During**	**After**	**t**	**P**	**t**	**P**
Ano-genital inspection	44	0.07 [0.04]	0.80 [0.10]	0.07 [0.04]	8.32	**<0.0001**	8.32	**<0.0001**
Copulation	14	0.00 [0.00]	1.21 [0.21]	0.00 [0.00]	6.96	**<0.0001**	6.96	**<0.0001**
Give grooming	354	0.06 [0.01]	0.32 [0.03]	0.04 [0.01]	10.50	**<0.0001**	11.18	**<0.0001**
Receive grooming	374	0.07 [0.01]	0.34 [0.03]	0.03 [0.01]	10.56	**<0.0001**	12.26	**<0.0001**
Handle infant	11	0.09 [0.09]	1.00 [0.13]	0.00 [0.00]	7.11	**<0.0001**	7.82	**<0.0001**
Give lip smack	129	0.02 [0.01]	0.28 [0.04]	0.05 [0.02]	6.98	**<0.0001**	6.16	**<0.0001**
Receive lip smack	14	0.00 [0.00]	0.07 [0.07]	0.07 [0.07]	0.87	0.3918	0.00	1.0000

Table 6.4. Results of Linear Mixed Models comparing sequence size and complexity measures of sequences produced outside of affiliative behaviors (sequence type: 'non-social') and during one of the affiliative behaviors.

Sequence type	**N**	**Sequence size**			**Derived call repertoire**			**# Exhaled moans**			**# Wobbles**			**# Yawns**		
		mean [SEM]	**t**	**p**	**mean [SEM]**	**t**	**p**	**mean [SEM]**	**t**	**p**	**mean [SEM]**	**t**	**p**	**mean [SEM]**	**t**	**p**
Non-social	1003	4.02 [0.11]	-	-	0.29 [0.02]	-	-	0.24 [0.01]	-	-	0.02 [0.01]	-	-	0.05 [0.01]	-	-
Give approach	180	6.24 [0.33]	7.40	**<0.0001**	0.57 [0.05]	5.02	**<0.0001**	0.37 [0.04]	2.49	**0.0130**	0.15 [0.05]	4.77	**<0.0001**	0.17 [0.03]	4.86	**<0.0001**
Receive approach	78	5.45 [0.47]	4.53	**<0.0001**	0.49 [0.08]	3.22	**0.0013**	0.31 [0.06]	1.60	0.1096	0.17 [0.06]	5.61	**<0.0001**	0.13 [0.04]	0.99	0.3223
Ano-genital inspection	35	6.03 [0.71]	2.73	**0.0065**	0.71 [0.12]	3.54	**0.0004**	0.37 [0.09]	0.43	0.6641	0.40 [0.13]	9.85	**<0.0001**	0.14 [0.06]	1.31	0.1915
Copulation	17	7.35 [0.92]	3.47	**0.0006**	0.65 [0.15]	1.83	0.0669	0.29 [0.11]	-0.47	0.6356	0.65 [0.26]	11.66	**<0.0001**	0.06 [0.06]	0.10	0.9195
Give grooming	112	5.50 [0.41]	4.44	**<0.0001**	0.44 [0.06]	0.98	0.3263	0.30 [0.05]	0.25	0.8039	0.24 [0.07]	6.83	**<0.0001**	0.04 [0.02]	-3.12	**0.0019**
Receive grooming	127	4.46 [0.31]	1.66	0.0963	0.57 [0.07]	5.25	**<0.0001**	0.31 [0.05]	1.87	0.0611	0.06 [0.02]	2.56	**0.0105**	0.24 [0.04]	6.62	**<0.0001**
Handle infant	11	6.00 [1.26]	2.79	**0.0053**	0.55 [0.21]	1.94	0.0532	0.18 [0.12]	0.23	0.8147	0.09 [0.09]	1.06	0.2899	0.27 [0.14]	2.25	**0.0247**
Give lip smack	36	5.72 [0.67]	2.30	**0.0218**	0.58 [0.11]	2.63	**0.0087**	0.33 [0.11]	0.84	0.4038	0.25 [0.09]	6.33	**<0.0001**	0.14 [0.06]	1.09	0.2746

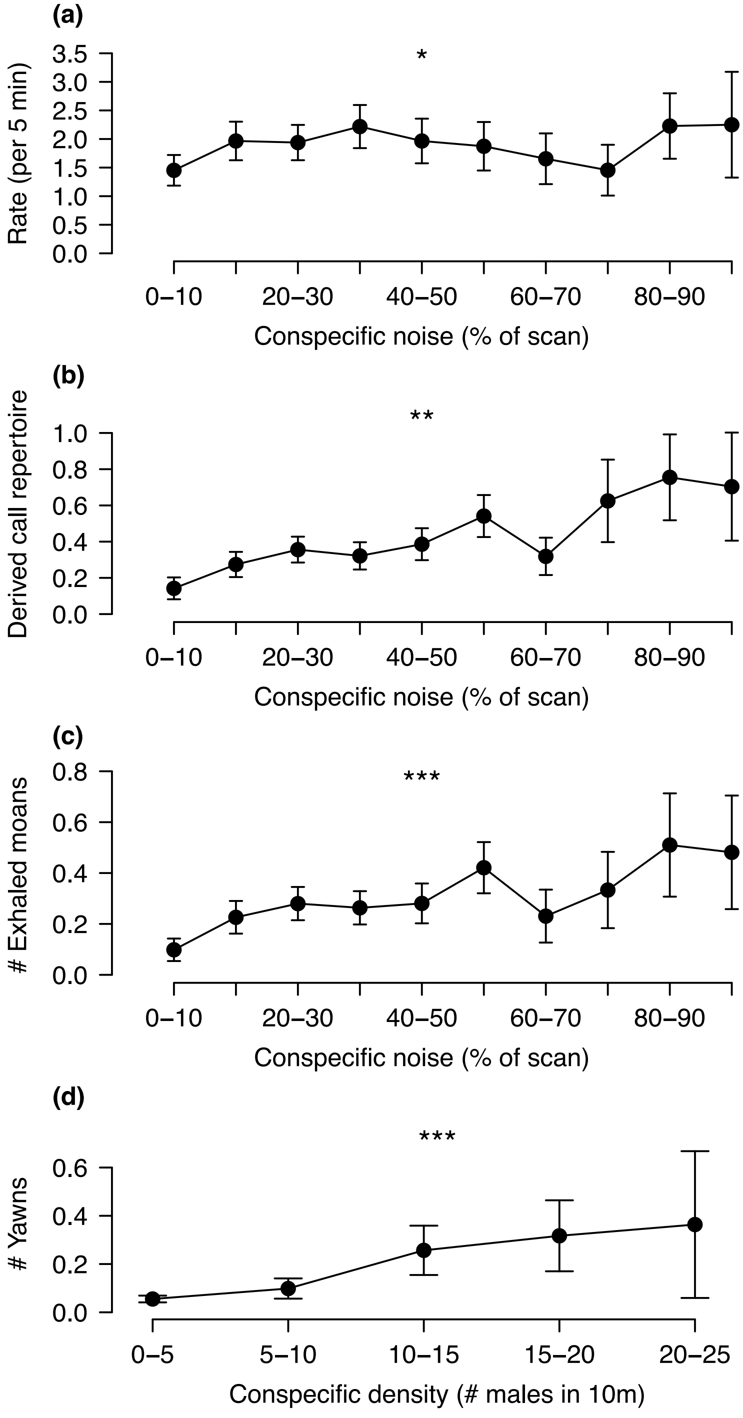


Figure 6.1. Relationships between vocal sequence production and complexity and features of the social environment. (a) Sequence rate, (b) number of derived call types (exhaled moans, wobbles and yawns) per sequence, and (c) number of exhaled moans per sequence were positively associated with conspecific noise levels. (d) Sequences contained a yawn if conspecific density was higher. Points and whiskers represent means plus two SEM. *p < 0.05, **p < 0.01, ***p < 0.001.

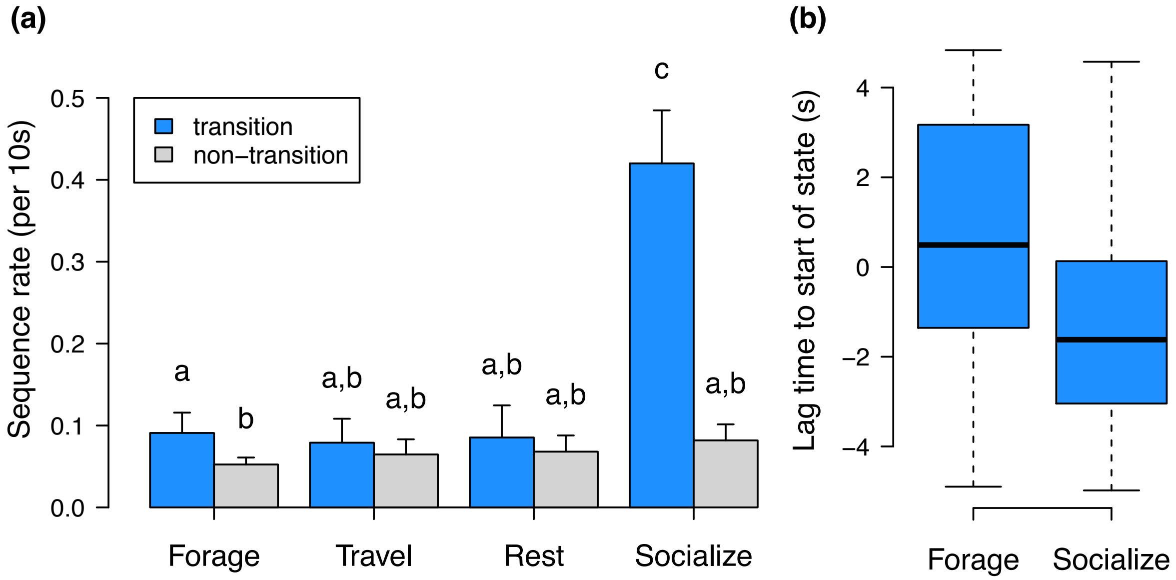
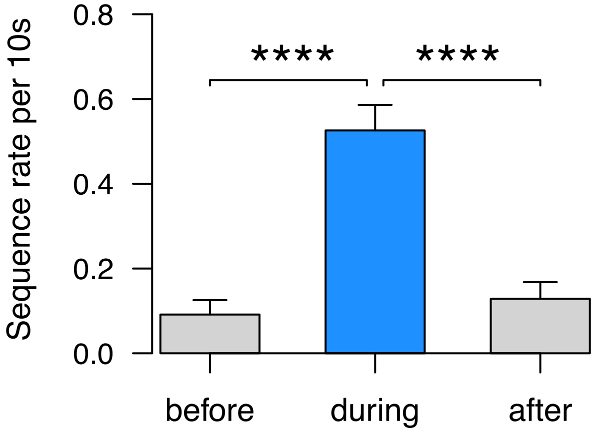


Figure 6.2. Sequence production during the transition and non-transition periods of different behavioral states. Pairwise comparisons were used to test (a) differences in sequence rates in transition and non-transition periods of behavioral states. Bars and whiskers represent means plus two SEM, and rates that are different from one another do not share any letters ($p < 0.05$). (b) Boxplots (with whiskers representing 95% confidence intervals) showing the latencies between the start of a sequence relative to the start of a behavioral state.

(a) Give Approach



(b) Receive Approach

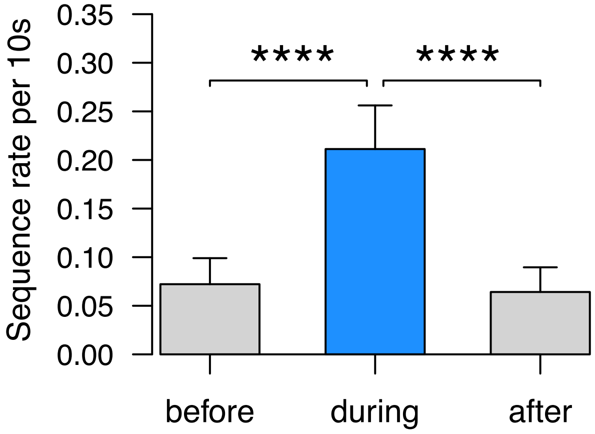


Figure 6.3. Comparisons of sequence production in the 10 s before, 10 s during, and 10 s after males (a) gave or (b) received an approach. Bars and whiskers represent means plus two SEM. **** $p < 0.0001$

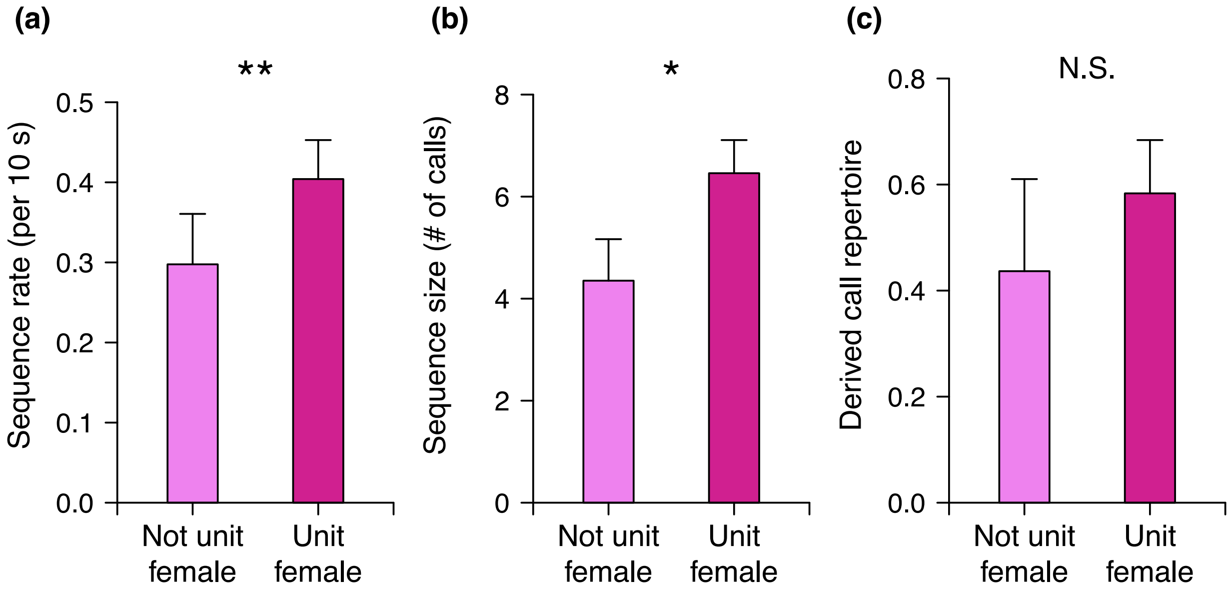


Figure 6.4. Comparisons in sequence production and complexity during approaches with adult unit females or other individuals. Bars and whiskers represent means plus two SEM. * $p < 0.05$, ** $p < 0.01$, $p > 0.05$ denoted as N.S.

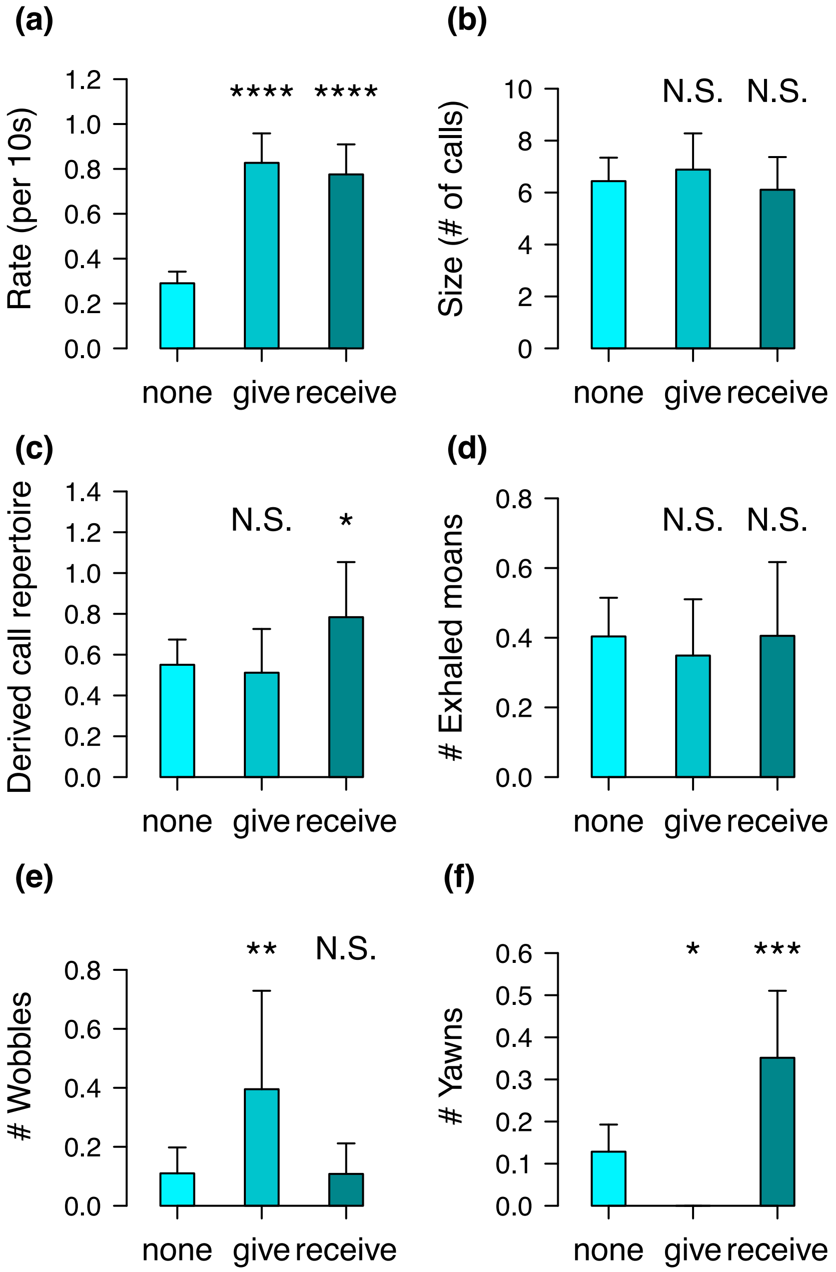


Figure 6.5. Grooming outcomes based on the production (a-b) and complexity (c-f) of sequences made during males' approaches to unit females. Comparisons were made between a no grooming outcome and outcomes that involved males either giving or receiving grooming with a unit female. Bars and whiskers represent means plus two SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, $p > 0.05$ denoted as N.S.

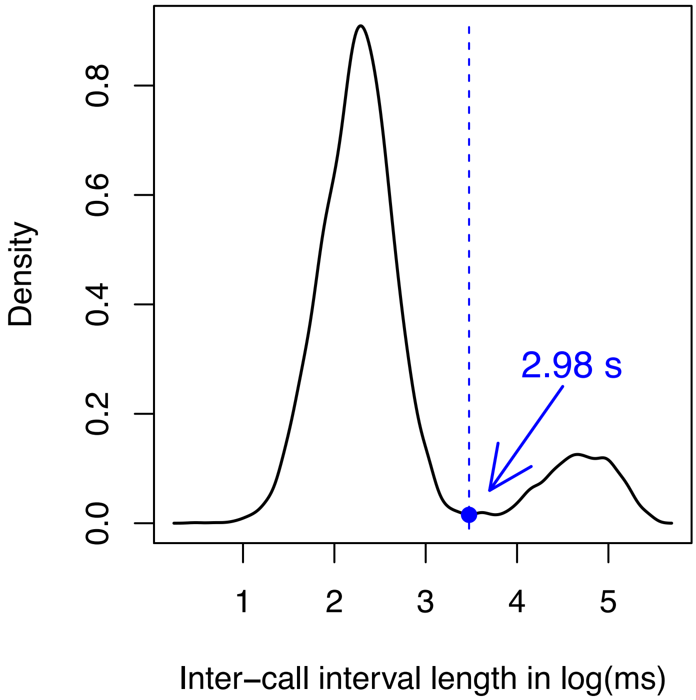


Figure 6.S1. Density plots of male inter-call intervals from continuous 3-5 minute acoustic recordings. The blue dotted line and point indicates the lowest point between the adjacent peaks and the cut-off thresholds for vocal sequence boundaries.

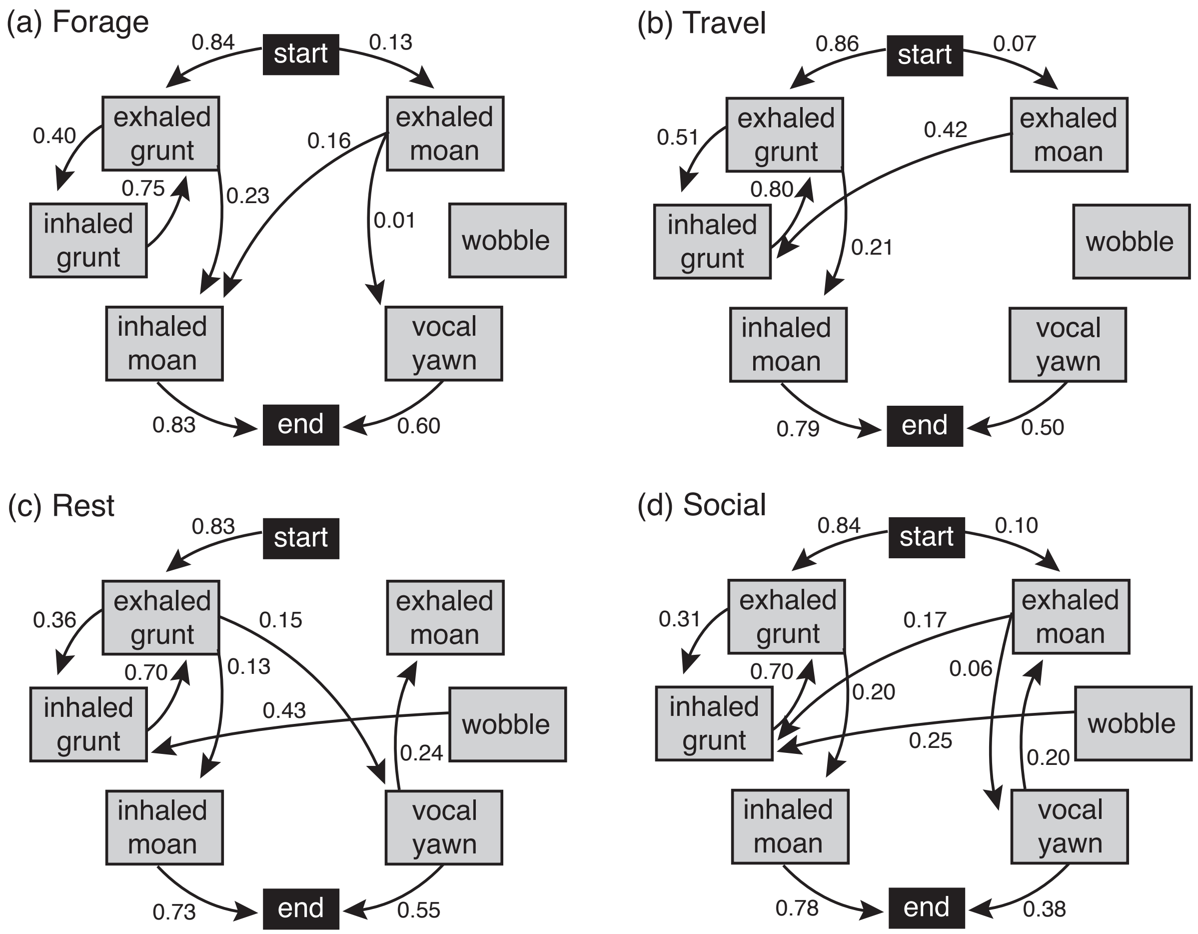


Figure 6.S2. Flow diagrams showing conditional probabilities in gelada male vocal sequences in the four behavioral states (a) foraging, (b) traveling (c) resting and (d) socializing. Arrows represent the conditional probability of occurrence (# of transitions of call pair divided by the total number of transitions starting with the initial call type). Arrow thicknesses are proportional to the corresponding transitional probability, which is written next to each arrow. All transitional probabilities shown are significant at $p < 0.05$.

Chapter 7:

Male geladas use complex vocal sequences to manage conflicts

Abstract

Vertebrates engage in a wide range of behavioral strategies to manage conflict, and the majority of these strategies are tactile in form (e.g., allogrooming). However, little is known about how animals utilize non-tactile behaviors, such as vocalizations, to manage conflict. Here, we study the post conflict behavior of wild geladas (*Theropithecus gelada*), a species known for producing vocal sequences composed of a diverse repertoire of affiliative call types. We collected male and female focal samples during control and post-conflict periods, as well as during 3 minute post-conflict (PC) and matched-control (MC) videos from male geladas. Conflicts consisted of threats, chases and physical aggression within and between harem-like reproductive units, as well as with bachelor groups. Both males and females engaged in non-aggressive approaches and agonistic interactions after conflicts, and males engaged in self-directed behaviors after conflicts. The sexes differed in their grooming and vocal responses to conflict: females engaged in grooming while males engaged in vocalizations. In addition, males produced vocal sequences sooner in PC sessions than MC sessions and these post-conflict sequences were longer and contained more derived call types. These findings suggest that gelada males preferentially use complex vocal sequences to engage in conciliation or reconciliation with multiple partners. We

outline the implications of these findings for understanding the role of social complexity (e.g., group size) in promoting vocal complexity.

Introduction

Species gain several benefits from being social, including protection from predators, resource defense, and shared knowledge (Alexander, 1974). One of the primary costs of sociality, however, are the conflicts resulting from competition over resources, which can lead increased physiological stress, injury, and even death (Smith, 1974). As a solution to this problem, many species engage in affiliative behavioral strategies after the occurrence of conflicts, and such strategies are thought to serve reconciliatory, conciliatory, and consolation functions (Aureli, 2002; Silk, 1998). Post-conflict behavioral strategies typically manifest in the form of body touching and allogrooming and are found in a wide range of taxa, including bottlenose dolphins (*Tursiops Truncatus*), corvid species, domestic goats (*Capra hircus*), horses (*Equus caballus*), spotted hyenas (*Crocuta crocuta*), dogs and wolves (*Canis familiaris* and *lupus*), wallabies (*Macropus rufogriseus*), and non-human primates (Baan, Bergmüller, Smith, & Molnar, 2014; Cools, Van Hout, & Nelissen, 2008; Cordoni & Norscia, 2014; Cozzi, Sighieri, Gazzano, Nicol, & Baragli, 2010; Fraser & Bugnyar, 2011; Logan, Emery, & Clayton, 2013; Silk, 2002; Tamaki, Morisaka, & Taki, 2006; Wahaj, Guse, & Holekamp, 2001). Interestingly, the primary feature in common with the majority of affiliative post-conflict strategies studied to date is that they are tactile in form (McFarland & Majolo, 2013).

An alternative to tactile-based strategies is to engage in vocal communication, a complex conflict management strategy used predominantly by humans (Aureli & de Waal, 2000; Blum-kulka & Olshtain, 1984; de Waal, 2000; Silk, 1998). There are many potential benefits of using vocal over tactile post-conflict strategies. Vocalizations can reach several recipients at once, thereby serving reconciliatory and conciliatory functions simultaneously. Vocalizations are also ‘cheap’ in that they require little energy to produce and can easily co-occur with other behaviors (Dunbar, 1998, 2003; Silk, Kaldor, & Boyd, 2000). Despite the potential to be an incredibly effective tool to manage conflicts, empirical evidence for vocal strategies among non-human taxa is surprisingly rare. The only systematic evidence of vocal post-conflict strategies comes from work on the soft tonal ‘grunt’ calls of chacma baboons (*Papio ursinus*) (Cheney & Seyfarth, 1997; Cheney, Seyfarth, & Silk, 1995; Silk, Cheney, & Seyfarth, 1996; Wittig, Crockford, Wikberg, Seyfarth, & Cheney, 2007). Baboons groom each other following conflicts just like several other non-human primates. However, dominant female baboons will sometimes direct grunts towards subordinates following fights, even in the absence of grooming.

Although baboons are the only non-human taxa shown to engage in a vocal post-conflict strategy, comparative work suggests that this strategy should be more common, at least in non-human primates. Primate taxa living in larger social groups, where more opportunities for conflict exist, produce larger repertoires of vocalizations (McComb & Semple, 2005). It has been an ongoing puzzle as to why these more diversified forms of vocal communication are needed in larger social groups (Freeberg, 2006; Freeberg, Ord, & Dunbar, 2012; Gustison, le Roux, & Bergman, 2012; Wilkinson, 2003). One hypothesis, the “vocal grooming hypothesis”, states that individuals have a limited amount of time to invest in socializing and therefore need more efficient strategies to maintain social bonds with several partners (Dunbar, 1998, 2003). In contexts when the stability of multiple social bonds is at risk, therefore, it would be expected that vocalizations are favored over one-on-one forms of affiliative interaction like grooming.

To test the hypothesis that species living in large and highly social groups favor vocal communication as a post-conflict strategy over tactile strategies, we investigated the post-conflict behavior of wild geladas (*Theropithecus gelada*), a primate known for living in extremely large communities (up to 1,200 individuals) and for engaging in high levels of both non-vocal and vocal affiliative behaviors (Aich, Moos-Heilen, & Zimmermann, 1990; Richman, 1976, 1987; Snyder-Mackler, Beehner, & Bergman, 2012; Tinsley Johnson, Snyder-Mackler, Beehner, & Bergman, 2014). These larger communities break down into smaller harem-like reproductive units comprised of one dominant leader male, up to a few follower males, several related females and their dependent offspring. Thus, conflicts can occur at several levels in gelada society – within units, between units, and between units and all-male bachelor groups. Preliminary findings from captive studies of post-conflict behavior suggest that geladas engage in vocal and non-vocal affiliative interactions after conflicts (Leone & Palagi, 2010; Swedell, 1997). Due to limitations in sample size and the types of conflicts available in captive studies, it is still unknown whether specific types of post-conflict behaviors characterize geladas (Leone & Palagi, 2010).

Compared to other Papionins, wild geladas vocalize at higher rates and spend large amounts of time grooming (Aiello & Dunbar, 1993; Gustison et al., 2012; McComb

& Semple, 2005). Moreover, wild geladas have a more diversified vocal repertoire than their baboon relatives, with males producing several species-specific call types: inhaled grunts, inhaled moans, exhaled moans, wobbles and vocalized yawns (Gustison et al., 2012). All the unique, or 'derived', call types are combined with exhaled grunts to produce complex vocal sequences. Sequences containing the three most acoustically salient derived calls – exhaled moans, wobbles and yawns – yield greater responses from females (Gustison & Bergman, n.d.-a, 2016). These complex vocal sequences facilitate social interaction between males and females, and there is preliminary evidence that they may even serve consolation functions following female-female conflicts (Gustison & Bergman, n.d.-b; Gustison et al., 2012).

We studied three aspects wild gelada post-conflict behavior. First, we characterized whether there were any specific types of non-vocal behaviors (e.g., approaches and grooming) that characterized gelada male or female post-conflict strategies. This included testing whether any features of conflicts (e.g., unit membership of opponents, intensity of the conflict) predicted non-vocal strategies. Second, we tested whether males or females responded to conflict with vocal strategies and whether a vocal strategy was predicted by any features of the conflict. For males, we also tested whether 'complexity' measures (e.g., sequence size and number of derived call types) specifically characterized post-conflict vocal behavior. Finally, we tested whether post-conflict vocal production or complexity impacted the subsequent affiliative behavior of receivers.

Methods

Study site and subjects

Data for this study come from adult geladas across three different bands in one wild community (about 1,200 individuals) living in the Sankaber area of the Simien Mountains National Park, Ethiopia (Snyder-Mackler et al., 2012). Gelada bands are comprised of smaller harem-like units with a leader male, 0-3 subordinate follower males, and 1-11 females with their immature offspring (Snyder-Mackler et al., 2012). This population has been under intensive behavioral study since January 2006, and study subjects are fully habituated to human observers on foot (approach distances less than 3 m) (Bergman, 2010). The gelada research was approved by the University Committee on Use and Care of Animals at the University of Michigan and adhered to the laws and guidelines of Ethiopia.

Focal sample collection

From January 2009 to June 2016, we collected 3,892 hr of 15 minute focal samples from 279 adult geladas. Focal samples included all-occurrence sampling (Altmann, 1974) of agonistic, affiliative, and self-directed behaviors on a palm pilot. Focal samples were categorized as ‘matched-control’ (MC) and ‘post-conflict’ (PC) following guidelines of previous studies (de Waal & Yoshihara, 1983). MC focal samples were those in which the focal animal did not express any agonistic behaviors in the first ten minutes of the sample, and PC focal samples were those in which the focal individual engaged in at least one agonistic behavior within the first 10 minutes. Agonistic behaviors included vocal threat grunts, visual threats (eye-brow raises or teeth grinding), chases, and physical aggression (biting, grabbing and hitting). Multiple agonistic events that occurred with the same opponent within a 10 s window were lumped together, and the end of the conflict was defined as the last agonistic event. PC focal samples were only used for analysis if the first agonistic behavior involved a single opponent for which the observer could determine the sex/age (immature individual, adult male or adult female), and unit membership (in the same reproductive unit, a different reproductive unit, or an all-male unit). We also categorized features of the conflicts, such as intensity and outcome. Conflict intensity was either mild (vocal/visual threats only) or severe (chases and/or physical aggression). Conflict outcome was whether the focal animal won the conflict, lost the conflict, or the winner could not be determined. A winner was decided if only one individual (the focal animal or the opponent) expressed a submissive vocalization (fear grunt or scream) or visual behavior (lip flip, crouch or presentation of the rump). To ensure that we had a representative dataset for all study animals, we retained focal samples for individuals with at least 10 MC and 10 PC focal samples. With these specifications in place, we retained 6,702 MC focal samples (13-249 samples per animal) and 967 PC focal samples (10-41 samples per animal) from 28 males and 26 females.

To test for potential post-conflict strategies, rates of affiliative, agonistic and self-directed behaviors were measured in the 10-13 min of MC focals and in the 3 min following the end of the first conflict in PC focals. Affiliative behaviors included ‘friendly’ approaches (did not occur simultaneously with agonistic behavior), copulation, grooming, lip-smacking, and contact call vocalizations (e.g., grunts). Agonistic behavior included any vocal/visual threats, chases or physical aggression. Self-directed behaviors included any scratching, self-grooming, body shakes or silent yawns. Finally, the behavioral state – foraging, traveling, resting, or socializing – at the start of the 3 min interval was recorded so that it could be included as a random effect in the analyses.

Video samples and acoustic analysis

One major limitation of using observational focal sampling to study post-conflict strategies is that it is impossible to explore the properties of vocal sequences that are characteristic of gelada males. As such, we supplemented our focal sample dataset on males by collecting a suite of video samples in control (MC) and post-conflict (PC) situations. This video-based PC-MC paradigm is based off established methods published by de Waal and others (de Waal & Yoshihara, 1983). From February 2013 to June 2014 we collected 400 3-minute video samples (MC = 200, PC = 200) from 20 gelada males (1-19 PC-MC pairs per male) using Flip UltraHD and Kodak Playsport video camcorders. PC video samples were collected by filming study males starting at the end of a conflict to at least 3 minutes afterwards. As with PC focal samples, the conflicts of PC video samples conflict could involve vocal/visual threats, chases or physical aggression. We also noted the opponents' sex/age and unit membership, conflict intensity and outcome, and the initial behavioral state of the male. The end of a conflict was defined as time at which the focal male ceased aggression and turned away from the opponent. We carried out MC video sessions by taking a 3-minute video recording of study males within two weeks after a PC session. The protocol for MC video recording was the same as PC session except that there were two conditions that had to be met: males could not have engaged in a conflict for at least 10 minutes prior to recording, and males had to start in the same type of behavioral state as the associated PC session (socializing/resting or foraging/traveling). All PC and MC videos were made between 0700 and 1730.

The acoustic data of each video file were extracted using Audacity (version 2.0.5) and saved as WAV files. Then, we used Praat (version 5.2.29) to locate vocal sequences and save each one as separate WAV files (see Fig 7.1 for examples). We define a vocal sequence as one or more discrete calls made by the same male and, for those sequences with two or more calls, separated by an inter-call interval of less than 2.983 s (Gustison & Bergman, n.d.-b). Vocal sequence files were given neutral numerical filenames so that they could be labelled by an observer blind to the caller identity and session type. The blind observer used Avisoft SAS-Lab Pro (version 5.1.12, Avisoft Bioacoustics) to generate spectrograms with a Fast Fourier transformation (size of 1,024 points) and labeled the start times, end times, and call types of all low-amplitude “contact calls” in a sequence. These contact calls included: exhaled grunts, inhaled grunts, exhaled moans, inhaled moans, wobbles and yawns (Fig 7.1) and are described in previous studies (Gustison & Bergman, 2016; Gustison, le Roux, & Bergman, 2012; Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016; Gustison et al. In Prep). We extracted 436 sequences (2,649 calls) from 179 PC videos and 224 sequences (979 calls) from 131 MC videos. The latency to give a vocal sequence was determined for each video (those without a sequence were given a 3 min latency). In addition, the size (total number of calls) and complexity of each sequence was determined following previous methods (Gustison & Bergman, n.d.-b). Sequence complexity was measured by derived call repertoire size and the total number of each derived call type per sequence. These derived calls included exhaled moans, wobbles and yawns. Finally, each video was scored for affiliative, agonistic, and self-directed behaviors so that latencies to non-vocal behaviors could be determined.

Data analysis

A series of Linear Mixed Models (LMM) were used to test the influence of conflicts on non-vocal and vocal behavior in geladas and the potential effect of post-conflict vocalizations on subsequent social interactions. First, we tested whether the rates of non-vocal affiliative, agonistic and self-directed behaviors differed in male and female MC and PC focal samples. In each LMM, the rate of a specific behavior type was the dependent variable, focal sample type (MC or PC) was a fixed effect with date, focal animal identity, and the initial behavioral state as random effects. Next, we explored whether the rates of non-vocal behaviors after a conflict depended on any features of the conflict (opponent unit membership, opponent sex/age, conflict intensity, and conflict outcome). These LMMs had behavior latencies as the dependent variable, conflict features as the fixed factors, and date, focal animal identity, and the initial behavioral state as random effects. Next, we explored whether male sequence production and complexity differed in MC and PC video sessions. We ran an LMM to test whether the vocal sequence rate, sequence size, derived call repertoire per sequence, number of exhaled moans, number of wobbles, or number of yawns per sequence differed in MC and PC videos. In all LMMs, video type (MC vs PC) was the fixed factor and male identity, date, and initial behavioral state (foraging/traveling or resting/socializing) were random effects. For LMMs involving sequence size and complexity, we restricted the male sequences to any produced during MC sessions (n = 390) and the first ones produced during PC sessions (n = 188).

We also explored whether PC male sequence production and complexity depended on any conflict features. These LMMs had sequence rate, latency, size or complexity measures as the dependent variable, conflict features as the fixed factors, and date, focal animal identity, and initial behavioral state as random effects. Last, we explored the social consequences of producing complex vocal sequences after conflicts. We first tested whether sequence production within the minute after conflict predicted the percent of time that male groomed or were groomed by females in their reproductive unit. In these LMMs, the percent of time spent grooming was the dependent variable, sequence production (yes or no) was the fixed factor, and male identity was the random effect. Then, we tested whether the grooming outcome (groomed by female or not) predicted the size and complexity of sequences produced within a minute after conflicts. In these LMMs, the sequence parameter (size, derived call repertoire, number of exhaled moans, number of wobbles, or number of yawns) was the dependent variable, grooming outcome was the fixed factor, and male identity was the random effect.

LMMs were run using the function “lmer” of the R package “lme4” (Bates, Maechler, & Bolker, 2012). The “lmerTest” package was implemented to determine the significance of the LMM coefficients (Kuznetsova, Brockhoff, & Christensen, 2013). We used the “difflsmeans” function in the lmerTest package to make pairwise comparisons. All calculations and statistical tests were carried out in R 3.3.0 (R Development Core Team, 2016). Critical values were set at alpha = 0.05, and all tests were two-tailed.

Results

Characterization of non-vocal behavior after conflicts

First, we used LMMs to test the effect of conflicts on non-vocal affiliative, agonistic and self-directed behaviors in male and female geladas (Table 7.1). We found that both sexes gave more friendly approaches after conflicts than during control periods. Also, females were more likely to give grooming after conflicts than during control periods. Neither sex showed differences in copulation or lip-smacking rates in MC and PC focal samples. Both sexes engaged in more agonistic behaviors after conflicts compared to control periods, and only males engaged in more self-directed behaviors after conflicts than during control periods. Then, we used LMMs to test whether rates of these behaviors after conflicts depended on features of the conflict: opponent unit membership, opponent age/sex, conflict intensity, and conflict outcome (Tables 7.2-7.3). There was some evidence that rates of approaches, grooming and aggression varied based on conflict features. Males were more likely to give a friendly approach if the opponent was a bachelor male than if the opponent was from the same reproductive unit. Females were more likely to groom after severe intensity conflicts compared to mild intensity conflicts. Males were less likely to engage in aggression after a conflict if they were the clear winner.

Characterization of vocal behavior after conflicts

We used LMMs to test whether vocal sequence rates in males and females differed in MC and PC focal samples. We found that male geladas gave vocal sequences at higher rates following conflicts than during control periods ($t = 6.88$, $p < 0.0001$; Fig 7.2a). However, female vocal rates did not differ across PC and MC focal samples ($t = 0.14$, $p = 0.8890$; Fig 7.2b). Next, we used LMMs to examine whether male sequence production and complexity differed across MC and PC video sessions. Males gave vocal sequences sooner after conflicts than during control sessions ($t = -13.60$, $p < 0.0001$ Fig 7.3a). Then, we used LMMs to compare the size and complexity of MC male sequences to the first sequences produced in PC sessions. Compared to MC sequences, the first PC sequences were longer ($t = 7.65$, $p < 0.0001$; Fig 7.3b) and contained a larger derived call repertoire ($t = 7.058$, $p < 0.0001$; Fig 7.3c), more exhaled moans ($t = 3.93$, $p = 0.0001$; Fig 7.3d), more wobbles ($t = 2.68$, $p = 0.0075$; Fig 7.3e), and more yawns ($t = 5.33$, $p < 0.0001$; Fig 7.3f).

Next, we tested whether male sequence production and complexity depended on conflict features (Tables 7.4-7.5). From the PC focal samples, we found that male sequence rate was higher if he was the winner of the conflict compared to when the conflict winner was undecided. From the PC video sessions, we found that males' latency to give a vocal sequence depended a variety of conflict characteristics. Males vocalized sooner if the conflict involved a same-unit opponent compared to a bachelor male. Males also vocalized sooner if the conflict opponent was another male rather than a female. Males vocalized sooner following a severe conflict compared to a mild conflict. Also, males vocalized sooner if they were the winner of a conflict rather than the loser of a conflict. Male vocal sequence size was larger if the opponent was from the same reproductive unit rather than a different reproductive unit. Sequence size was also larger if the opponent was another adult male rather than an immature individual. There was no evidence that the complexity of post-conflict sequences depended on conflict features.

Influence of post-conflict vocalizations on subsequent social interactions

We used LMMs to examine whether sequence production and complexity following a conflict influences prosocial interactions with unit females. First, we tested whether the production of a sequence in the first minute following a conflict (conflicts with sequences = 162, conflicts without sequences = 38) predicted whether males gave or received grooming with females from their unit. We did not find evidence that males gave different amounts of grooming after conflicts based on whether they produced a vocal sequence (t = 1.01, p = 0.3153; Fig 7.4a). However, males received more grooming following a conflict if they produced a sequence (t = 2.33, p = 0.0208; Fig 7.4b). Latencies between sequence production and grooming revealed that sequences occurred before grooming (Fig 7.4c). Males gave grooming about 44.62 ± 10.05 s after starting to vocalize (n = 29 videos) and received grooming about 55.56 ± 9.50 s after starting to vocalize (n = 31 videos). Next, we tested whether being groomed by a unit female predicted the size or complexity of sequences produced in the first minute after a conflict. We found no evidence that sequence size varied based on grooming outcome (t = 1.54, p = 0.1257; Fig 7.5a). However, sequences followed up with receipt of grooming contained larger derived call repertoire (t = 3.59, p = 0.0004; Fig 7.5b), more exhaled moans (t = 2.37, p = 0.0191; Fig 7.5c), more wobbles (t = 2.62, p = 0.0097; Fig 7.5d) and more yawns (t = 3.20, p = 0.0017; Fig 7.5e).

Discussion

Gelada males preferentially respond to conflicts with complex vocal sequences instead of other available tactile strategies like grooming. Currently, only a few non-human empirical studies show that vocalizations are used in reconciliation, conciliation or consolation contexts (Silk, 2002, 2007), and we are not aware of any that demonstrate that complex vocal strategies are preferred over tactile forms of interaction like ‘grooming’. Therefore, this study makes a significant contribution to our understanding of the evolution of conflict management strategies by showing evidence for a vocal strategy outside of humans. Specifically, our study supports the controversial ‘vocal grooming’ hypothesis that states that vocal strategies to maintain social bonds may outweigh one-on-one interactions in species living in comparatively large groups, like humans and geladas (Dunbar, 1998, 2003).

It has been an ongoing mystery why primates living in larger groups produce a more diversified repertoire of vocalizations (Bouchet, Blois-Heulin, & Lemasson, 2013; Gustison et al., 2012; McComb & Semple, 2005). The current findings show that the need for complex vocal strategies in specific contexts, like social instability, could be one of the driving forces in this association. Previous research indicates that longer, more complex vocal sequences function to facilitate male social bonds with female unit members (Gustison & Bergman, n.d.-b, 2016). Our study supports and expands on this interpretation by demonstrating that these complex calls may help males attenuate potential uncertainty in their social bonds with females during periods of instability. This interpretation is further supported by the finding that males giving more complex sequences after conflicts are also more likely to be subsequently groomed by females in their units.

The vocal post-conflict strategy was not generalizable to both sexes in geladas. Instead of vocalizing, females began to groom others after an approach. One explanation for this difference comes from the vocal grooming hypothesis in that male and female geladas differ in the number of social partners with which they maintain strong relationships. Although males maintain relationships with all females in their units, females tend to focus their socializing on specific individuals (Tinsley Johnson et al., 2014). Given this distinction, we would expect that males would prefer a vocal strategy in order to reach multiple social partners. On the other hand, it is unclear why female baboons, but not gelada females, engage in grunts after conflicts (Cheney & Seyfarth, 1997; Cheney et al., 1995; Silk et al., 1996). Future work will be needed to tease apart the potential situations in which gelada females engage in different post-conflict strategies.

One of the main questions asked when characterizing post-conflict behavior is whether opponents engage in reconciliatory or ‘triadic’ type conciliatory/consolation interactions (Fraser & Aureli, 2008; Koski & Sterck, 2007; Logan et al., 2013). While most studies of post-conflict behavior demonstrate that reconciliation occurs between opponents, some studies are starting to reveal that these strategies depend on the social system. For example, corvids are known to have the highest-quality relationships with their mates, and conflicts typically occur with individuals outside of these close relationships. Following a fight, corvids avoid former opponents and instead seek out triadic affiliative interactions with their mates (Fraser & Bugnyar, 2011; Logan et al., 2013). Our study suggests that gelada post-conflict behavior is more closely aligned with corvids over other non-human primates. Like corvids, most gelada conflicts occurred between individuals in different units, yet post-conflict behavior was relatively unaffected by the unit membership of opponents. Males almost always approached unit members while vocalizing, and this tended to be followed up by consolation-like behavior from unit females. A key difference between corvids and geladas is that gelada males specifically engage in vocal over tactile strategies. This means that a potential secondary benefit of these post-conflict vocalizations is that they affect non-unit members in addition to unit members. Future work will be needed to test whether gelada behavior is influenced by whether a non-unit opponent gave a post-conflict vocal sequence.

Currently, there is very little evidence to suggest that non-human species – other than humans, baboons and geladas – engage in affiliative vocal strategies to manage conflicts. It is important to note, however, that this lack of evidence does not mean that affiliative vocalizations are non-existent after conflicts. Studies of post-conflict behavior rarely include affiliative vocalizations in the ethograms, and so it is very possible that these affiliative vocalizations occur but are not measured. Victims of conflicts are often noted to produce loud ‘submissive’ vocalizations during conflicts (de Waal & van Roosmalen, 1979; Kutsukake & Clutton-Brock, 2008; Palagi, Antonacci, & Norscia, 2008). We suggest that future post-conflict studies also investigate the production of affiliative types of vocalization.

Acknowledgements

We thank the Ethiopian Wildlife Conservation Authority (EWCA) as well as the wardens and staff of the Simien Mountain National Park for permission and support in conduction research on geladas. We also thank all the members of the University of Michigan Gelada Research Project (UMGRP) for assistance in data collection and valuable insight on analyses and interpretation. MLG was supported by grants from the National Geographic Society (9122-12, W304-14), Leakey Foundation, University of Michigan, and Petridish (primary supporter: John Allen). UMGRP was supported by the Wildlife Conservation Society (SSF 67250), the National Geographic Society (8100-06), the Leakey Foundation, the National Science Foundation (BCS-0715179, BCS-0962118, IOS- 1255974), and the University of Michigan.

References

Aich, H., Moos-Heilen, R., & Zimmermann, E. (1990). Vocalizations of adult gelada baboons (*Theropithecus gelada*): Acoustic structure and behavioural context. *Folia Primatologica*, *55*, 109–132.

Aiello, L. C., & Dunbar, R. I. M. (1993). Neocortex size, group size, and the evolution of language. *Current Anthropology*, *34*, 184–193.

Alexander, R. D. (1974). The evolution of social behavior. *Annual Review of Ecology and Systematics*, *5*, 325–383.

Altmann, J. (1974). Observational study of behavior : Sampling methods. *Behaviour*, *49*, 227–267.

Aureli, F. (2002). Conflict resolution following aggression in gregarious animals: a predictive framework. *Animal Behaviour*, *64*, 325–343.

Aureli, F., & de Waal, F. B. M. (2000). Natural Conflict Resolution. In *Natural Conflict Resolution*. Berkeley, CA: California University Press.

Baan, C., Bergmüller, R., Smith, D. W., & Molnar, B. (2014). Conflict management in free-ranging wolves, *Canis lupus*. *Animal Behaviour*, *90*, 327–334.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. Retrieved from http://cran.r-project.org/ package=lme4

Bergman, T. J. (2010). Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 3045–53.

Blum-kulka, S., & Olshtain, E. (1984). Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics*, *5*, 196–213.

Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: a comparison of three non-human primate species. *Frontiers in Psychology*, *4*, 390.

Cheney, D. L., & Seyfarth, R. M. (1997). Reconciliatory grunts by dominant female baboons influence victims' behaviour. *Animal Behaviour*, *54*, 409–18.

Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Animal Behaviour*, *50*, 249–257.

Cools, A. K. a, Van Hout, A. J. M., & Nelissen, M. H. J. (2008). Canine reconciliation and third-party-initiated postconflict affiliation: Do peacemaking social mechanisms in dogs rival those of higher primates? *Ethology*, *114*, 53–63.

Cordoni, G., & Norscia, I. (2014). Peace-making in marsupials: The first study in the red-necked wallaby (*Macropus rufogriseus*). *PLoS ONE*, *9*, 1–7.

Cozzi, A., Sighieri, C., Gazzano, A., Nicol, C. J., & Baragli, P. (2010). Post-conflict friendly reunion in a permanent group of horses (*Equus caballus*). *Behavioural Processes*, *85*, 185–190.

de Waal, F. B. M. (2000). Primates--a natural heritage of conflict resolution. *Science*, *289*, 586–90.

de Waal, F. B. M., & van Roosmalen, A. (1979). Reconciliation and consolation among chimpanzees. *Behavioral Ecology and Sociobiology*, *5*, 55–66.

de Waal, F. B. M., & Yoshihara, D. (1983). Reconciliation and redirected affection in Rhesus monkeys. *Behaviour*, *85*, 224–241.

Dunbar, R. I. M. (1998). Grooming, Gossip, and the Evolution of Language. In *Grooming, Gossip, and the Evolution of Language*. Cambridge, MA: Harvard University Press.

Dunbar, R. I. M. (2003). The social brain: Mind, language, and society in evolutionary perspective. *Annual Reviews in Anthropology*, 32, 164-181.

Fraser, O. N., & Aureli, F. (2008). Reconciliation, consolation and postconflict behavioral specificity in chimpanzees. *American Journal of Primatology*, *70*, 1114–1123.

Fraser, O. N., & Bugnyar, T. (2011). Ravens reconcile after aggressive conflicts with valuable partners. *PLoS ONE*, *6*, e18118.

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

Freeberg, T. M., Ord, T. J., & Dunbar, R. I. M. (2012). The social network and communicative complexity: preface to theme issue. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1782–1784.

Gustison, M. L., & Bergman, T. J. (In Review: n.d.-a). Divergent acoustic properties of gelada and baboon vocalizations and their implications for the evolution of human speech. *Journal of Language Evolution*.

Gustison, M. L., & Bergman, T. J. (n.d.-b). The social functions of complex call sequences in wild geladas.

Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Scientific Reports*, *6*, 19680.

Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1847–59.

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*, E2750–E2758.

Koski, S. E., & Sterck, E. H. M. (2007). Triadic postconflict affiliation in captive chimpanzees: does consolation console? *Animal Behaviour*, *73*, 133–142.

Kutsukake, N., & Clutton-Brock, T. H. (2008). Do meerkats engage in conflict management following aggression? Reconciliation, submission and avoidance. *Animal Behaviour*, *75*, 1441–1453.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/package=lmerTest

Leone, A., & Palagi, E. (2010). Reconciling conflicts in a one-male society: The case of geladas (*Theropithecus gelada*). *Primates*, *51*, 203–212.

Logan, C. J., Emery, N. J., & Clayton, N. S. (2013). Alternative behavioral measures of postconflict affiliation. *Behavioral Ecology*, *24*, 98–112.

McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, *1*, 381–5.

McFarland, R., & Majolo, B. (2013). The importance of considering the behavioral form of reconciliation in studies of conflict resolution. *International Journal of Primatology*, *34*, 15–29.

Palagi, E., Antonacci, D., & Norscia, I. (2008). Peacemaking on treetops: first evidence of reconciliation from a wild prosimian (*Propithecus verreauxi*). *Animal Behaviour*, *76*, 737–747.

R Development Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Richman, B. (1976). Some vocal distinctive features used by gelada monkeys. *The Journal of the Acoustical Society of America*, *60*, 718–24.

Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates*, *28*, 199–223.

Silk, J. B. (1998). Making amends. Adaptative perspectives on conflict remediation in monkeys, apes and humans. *Human Nature*, *9*, 341–368.

Silk, J. B. (2002). The form and function of reconciliation in primates. *Annual Review of Anthropology*, *31*, 21–44.

Silk, J. B., Cheney, D. L., & Seyfarth, R. M. (1996). The form and function of post-conflict interactions between female baboons. *Animal Behaviour*, *52*, 259–268.

Silk, J. B., Kaldor, E., & Boyd, R. (2000). Cheap talk when interests conflict. *Animal Behaviour*, *59*, 423–432.

Smith, J. M. (1974). The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology*, *47*, 209–21.

Snyder-Mackler, N., Beehner, J. C., & Bergman, T. J. (2012). Defining higher levels in the multilevel societies of geladas (*Theropithecus gelada*). *International Journal of Primatology*, *33*, 1054–1068.

Swedell, L. (1997). Patterns of reconciliation among captive gelada baboons (*Theropithecus gelada*): A brief report. *Primates*, *38*, 325–330.

Tamaki, N., Morisaka, T., & Taki, M. (2006). Does body contact contribute towards repairing relationships? The association between flipper-rubbing and aggressive behavior in captive bottlenose dolphins. *Behavioural Processes*, *73*, 209–215.

Tinsley Johnson, E., Snyder-Mackler, N., Beehner, J. C., & Bergman, T. J. (2014). Kinship and Dominance Rank Influence the Strength of Social Bonds in Female Geladas (*Theropithecus gelada*). *International Journal of Primatology*, *35*, 288–304.

Wahaj, S. A., Guse, K. R., & Holekamp, K. E. (2001). Reconciliation in the spotted hyena (*Crocuta crocuta)*. *Ethology*, *107*, 1057–1074.

Wilkinson, G. S. (2003). Social and vocal complexity in bats. In F. B. M. de Waal & P. L.

Tyack (Eds.), *Animal social complexity: Intelligence, culture, and individualized societies* (pp. 322–341). Cambridge, MA: Harvard University Press.

Wittig, R. M., Crockford, C., Wikberg, E., Seyfarth, R. M., & Cheney, D. L. (2007). Kin-mediated reconciliation substitutes for direct reconciliation in female baboons. *Proceedings of the Royal Society, Series B: Biological Sciences*, *274*, 1109–1115.

Table 7.1. Comparisons of affiliative, agonistic and self-directed behavior rates across matched-control (MC) focals and post-conflict (PC) focals in males and females.

Behavior	**MC rate per 3 min**	**PC rate per 3 min**	**LMM results**	
	Mean [SEM]	**Mean [SEM]**	**t**	**p**
Males	**n = 2,222**	**n = 582**		
Friendly approach	0.10 [0.01]	0.25 [0.02]	**7.43**	**<0.0001**
Copulation	0.01 [0.00]	0.02 [0.01]	0.96	0.3363
Give grooming	0.18 [0.01]	0.13 [0.02]	0.15	0.8835
Lip-smacking	0.05 [0.01]	0.04 [0.01]	0.57	0.5670
Aggression	0.07 [0.01]	0.37 [0.03]	**14.93**	**<0.0001**
Self-directed behavior	0.33 [0.02]	0.49 [0.04]	**3.03**	**0.0025**
Females	**n = 4,480**	**n = 385**		
Friendly approach	0.10 [0.01]	0.18 [0.02]	**3.09**	**0.0002**
Copulation	0.00 [0.00]	0.00 [0.00]	-0.05	0.9607
Give grooming	0.20 [0.01]	0.23 [0.03]	**3.14**	**0.0017**
Lip-smacking	0.02 [0.00]	0.01 [0.01]	-0.84	0.4035
Aggression	0.03 [0.00]	0.13 [0.02]	7.68	**<0.0001**
Self-directed behavior	0.12 [0.01]	0.12 [0.02]	-0.15	0.8804

Table 7.2. Mean rates of affiliative, agonistic and self-directed behaviors in the 3 minutes following different types of conflicts that occurred during focal samples.

Behavior	Opponent unit			Opponent sex/age			Conflict intensity		Conflict outcome		
	Same unit	**Different unit**	**All-male unit**	**Same-sex**	**Opposite-sex**	**Immature**	**Mild**	**Severe**	**Winner**	**Loser**	**Unknown**
Males	**n = 136**	**n = 416**	**n = 30**	**n = 330**	**n = 175**	**n = 77**	**n = 521**	**n = 61**	**n = 132**	**n = 9**	**n = 441**
Friendly approach	0.26 [3.92]	0.24 [0.03]	0.50 [0.13]	0.24 [0.03]	0.25 [0.04]	0.32 [0.07]	0.25 [0.02]	0.28 [0.07]	0.24 [0.05]	0.22 [0.15]	0.26 [0.03]
Copulation	0.00 [0.00]	0.02 [0.01]	0.03 [0.03]	0.02 [0.01]	0.02 [0.01]	0.00 [0.00]	0.02 [0.01]	0.02 [0.02]	0.02 [0.01]	0.00 [0]	0.02 [0.01]
Give grooming	0.15 [0.03]	0.11 [0.02]	0.17 [0.11]	0.10 [0.02]	0.17 [0.04]	0.14 [0.05]	0.12 [0.02]	0.13 [0.04]	0.15 [0.04]	0.44 [0.34]	0.11 [0.02]
Lip-smacking	0.05 [0.02]	0.03 [0.01]	0.00 [0.00]	0.03 [0.01]	0.07 [0.02]	0.00 [0.00]	0.04 [0.01]	0.03 [0.02]	0.05 [0.02]	0.11 [0.11]	0.03 [0.01]
Aggression	0.43 [0.07]	0.31 [0.03]	0.77 [0.21]	0.35 [0.04]	0.41 [0.07]	0.34 [0.08]	0.35 [0.03]	0.51 [0.15]	0.31 [0.05]	1.33 [0.67]	0.36 [0.03]
Self-directed	0.36 [0.07]	0.50 [0.05]	0.83 [0.22]	0.49 [0.06]	0.55 [0.08]	0.31 [0.08]	0.50 [0.05]	0.34 [0.09]	0.35 [0.07]	0.56 [0.38]	0.53 [0.05]
Females	**n = 143**	**n = 241**	**n = 1**	**n= 243**	**n = 37**	**n = 105**	**n = 309**	**n = 76**	**n = 78**	**n = 14**	**n = 293**
Friendly approach	0.17 [3.92]	0.19 [0.03]	0	0.21 [0.03]	0.19 [0.07]	0.12 [0.04]	0.17 [0.03]	0.22 [0.05]	0.17 [0.05]	0.14 [0.1]	0.19 [0.03]
Copulation	0	0	0	0	0	0	0	0	0	0	0
Give grooming	0.34 [0.06]	0.17 [0.03]	0	0.21 [0.03]	0.16 [0.06]	0.32 [0.06]	0.17 [0.03]	0.50 [0.08]	0.36 [0.09]	0.14 [0.10]	0.20 [0.03]
Lip-smacking	0.01 [0.01]	0.01 [0.01]	0	0.01 [0.01]	0.03 [0.03]	0.00 [0.00]	0.01 [0.00]	0.03 [0.02]	0.01 [0.01]	0.07 [0.07]	0.01 [0.00]
Aggression	0.11 [0.03]	0.14 [0.02]	0	0.15 [0.02]	0.03 [0.03]	0.12 [0.04]	0.13 [0.02]	0.13 [0.04]	0.09 [0.03]	0.07 [0.07]	0.14 [0.02]
Self-directed	0.06 [0.02]	0.15 [0.03]	0	0.14 [0.03]	0.16 [0.07]	0.04 [0.02]	0.12 [0.03]	0.12 [0.04]	0.09 [0.04]	0.00 [0.00]	0.13 [0.03]

Table 7.3. Results of Linear Mixed Models comparing rates of affiliative, agonistic and self-directed behaviors in the 3 minutes following different types of conflicts that occurred during focal samples.

	Opponent unit				**Opponent sex/age**				**Conflict intensity**		**Conflict outcome**			
Behavior latency	**Same vs Different**		**Same vs. All-male**		**Same sex vs Opposite**		**Same sex vs Immature**		**Mild vs. Severe**		**Winner vs. Loser**		**Winner vs. Undecided**	
	t	**p**	**t**	**p**	**t**	**p**	**t**	**p**	**t**	**p**	**t**	**p**	**t**	**p**
Males														
Friendly approach	0.01	0.9881	**2.03**	**0.0431**	0.58	0.5634	1.27	0.2054	0.24	0.8132	-0.18	0.8540	0.50	0.6204
Copulation	1.81	0.0701	1.68	0.0926	0.94	0.3499	-0.29	0.7752	0.22	0.8239	-0.47	0.6393	-0.79	0.4317
Give grooming	0.25	0.8043	0.82	0.4136	1.17	0.2441	0.59	0.5533	-0.96	0.3373	1.59	0.1114	-0.96	0.3394
Lip-smacking	-0.80	0.4248	-1.18	0.2376	1.04	0.2986	-1.50	0.1354	-0.95	0.3407	0.54	0.5870	-0.72	0.4699
Aggression	-1.56	0.1185	1.83	0.0672	1.42	0.1560	-0.30	0.7666	1.80	0.0730	**4.28**	**<0.0001**	**2.00**	**0.0455**
Self-directed	0.32	0.7506	1.46	0.1437	1.92	0.0550	-0.10	0.9187	-0.62	0.5365	0.61	0.5410	1.50	0.1347
Females														
Friendly approach	0.59	0.5587	-0.38	0.7079	-0.61	0.5442	-1.93	0.0548	1.32	0.1886	-0.38	0.7060	0.01	0.9937
Copulation	0.29	0.7736	0.04	0.9685	-0.40	0.6886	-0.31	0.7583	-0.10	0.9229	0.09	0.9296	0.21	0.8362
Give grooming	-0.06	0.9519	0.05	0.9606	-1.13	0.2590	-1.02	0.3078	**2.60**	**0.0096**	-0.98	0.3300	-1.50	0.1344
Lip-smacking	-0.28	0.7793	-0.10	0.9171	-0.09	0.9248	-1.49	0.1365	1.15	0.2508	1.30	0.1958	-0.37	0.7143
Aggression	0.11	0.9149	-0.14	0.8887	-1.90	0.0584	-0.28	0.7799	0.71	0.4797	0.40	0.6917	1.11	0.2693
Self-directed	1.00	0.3199	-0.31	0.7573	1.02	0.3084	-1.60	0.1102	1.66	0.0981	-1.21	0.2291	0.43	0.6709

Table 7.4. Vocal sequence production and complexity in the 3 minutes following different types of conflicts that occurred during focal samples. Results listed as means [SEM].

Behavior	**Opponent unit**			**Opponent sex/age**			**Conflict intensity**		**Conflict outcome**		
	Same unit	**Different unit**	**All-male unit**	**Same-sex**	**Opposite-sex**	**Immature**	**Mild**	**Severe**	**Winner**	**Loser**	**Unknown**
Female focal samples	**n = 143**	**n = 241**	**n = 1**	**n = 243**	**n = 37**	**n = 105**	**n = 309**	**n = 76**	**n = 78**	**n = 14**	**n = 293**
Sequence rate (per 3 min)	0.41 [0.06]	0.68 [0.06]	0	0.64 [0.06]	0.57 [0.12]	0.44 [0.08]	0.63 [0.05]	0.38 [0.08]	0.38 [0.08]	0.43 [0.17]	0.64 [0.05]
Male focal samples	**n = 136**	**n = 416**	**n = 30**	**n = 330**	**n = 175**	**n = 77**	**n = 521**	**n = 61**	**n = 132**	**n = 9**	**n = 441**
Sequence rate (per 3 min)	1.22 [0.11]	1.12 [0.06]	1.23 [0.22]	1.13 [0.06]	1.17 [0.09]	1.17 [0.17]	1.16 [0.06]	1.05 [0.12]	1.41 [0.12]	1.33 [0.44]	1.07 [0.06]
Male video sessions	**n = 49**	**n = 116**	**n = 35**	**n = 144**	**n = 37**	**n = 19**	**n = 111**	**n = 89**	**n = 51**	**n = 8**	**n = 141**
Latency to vocalize (s)	19.51 [6.6]	36.82 [5.34]	44.57 [11.04]	34.98 [4.67]	30.65 [10.21]	32.44 [12.65]	42.78 [5.98]	22.91 [4.89]	17.82 [5.67]	62.79 [26.85]	38.13 [5.02]
Male video sequences	**n = 47**	**n = 108**	**n = 33**	**n = 138**	**n = 33**	**n = 17**	**n = 103**	**n = 85**	**n = 49**	**n = 6**	**n = 133**
Sequence size	9.13 [0.77]	6.17 [0.43]	8.58 [1.32]	7.22 [0.48]	8.55 [0.98]	5.88 [0.89]	6.38 [0.47]	8.48 [0.66]	8.1 [0.76]	3.83 [0.60]	7.20 [0.49]
Derived call repertoire	0.79 [0.1]	0.56 [0.06]	0.85 [0.15]	0.64 [0.06]	0.88 [0.12]	0.53 [0.12]	0.61 [0.07]	0.74 [0.08]	0.78 [0.09]	0.33 [0.21]	0.65 [0.06]
# of exhaled moans	0.53 [0.1]	0.44 [0.05]	0.33 [0.09]	0.43 [0.05]	0.42 [0.1]	0.59 [0.15]	0.50 [0.06]	0.38 [0.06]	0.51 [0.09]	0.17 [0.17]	0.43 [0.05]
# of wobbles	0.21 [0.09]	0.08 [0.03]	0.24 [0.12]	0.12 [0.04]	0.33 [0.12]	0.00 [0.00]	0.06 [0.03]	0.25 [0.07]	0.22 [0.08]	0.00 [0.00]	0.12 [0.04]
# of yawns	0.19 [0.06]	0.08 [0.03]	0.39 [0.09]	0.17 [0.03]	0.24 [0.08]	0.00 [0.00]	0.11 [0.03]	0.24 [0.05]	0.16 [0.05]	0.17 [0.17]	0.17 [0.03]

Table 7.5. Results of Linear Mixed Models comparing vocal sequence production and complexity in the 3 minutes following different types of conflicts that occurred during focal samples.

	Opponent unit				**Opponent sex/age**				**Conflict intensity**		**Conflict outcome**			
Measure	**Same vs Different**		**Same vs. All-male**		**Same sex vs Opposite**		**Same sex vs Immature**		**Mild vs. Severe**		**Winner vs. Loser**		**Winner vs. Undecided**	
	t	**p**	**t**	**p**	**t**	**p**	**t**	**p**	**t**	**p**	**t**	**p**	**t**	**p**
Vocal sequence rate & latency														
Female rate	0.89	0.3733	-0.59	0.5534	0.15	0.8805	-0.06	0.9532	-1.08	0.2814	0.48	0.6321	0.90	0.3685
Male rate	0.67	0.5010	0.13	0.8934	0.27	0.7903	-0.42	0.6744	-1.42	0.1562	-0.70	0.4844	**-3.23**	**0.0013**
Male latency	0.96	0.3375	**2.00**	**0.0471**	**2.31**	**0.0217**	1.77	0.0789	**-2.43**	**0.0159**	**2.66**	**0.0084**	1.72	0.0876
Male sequence characteristics														
Sequence size	**-3.15**	**0.0019**	-1.43	0.1541	-0.66	0.5073	-2.01	**0.0457**	1.62	0.1066	-0.90	0.3695	0.79	0.4305
Derived call repertoire	-0.78	0.4349	0.48	0.6346	1.04	0.3012	-0.93	0.3526	0.04	0.9652	-0.61	0.5449	0.18	0.8553
# of exhaled moans	-1.52	0.1296	-1.45	0.1501	-0.92	0.3592	-0.31	0.7577	-1.52	0.1297	-1.14	0.2540	-0.14	0.8866
# of wobbles	0.05	0.9577	0.78	0.4370	1.32	0.1879	-0.75	0.4536	1.74	0.0831	-0.13	0.8931	-0.04	0.9679
# of yawns	-0.71	0.4794	1.96	0.0515	1.25	0.2133	-1.02	0.3069	1.17	0.2449	0.52	0.6031	0.38	0.7053

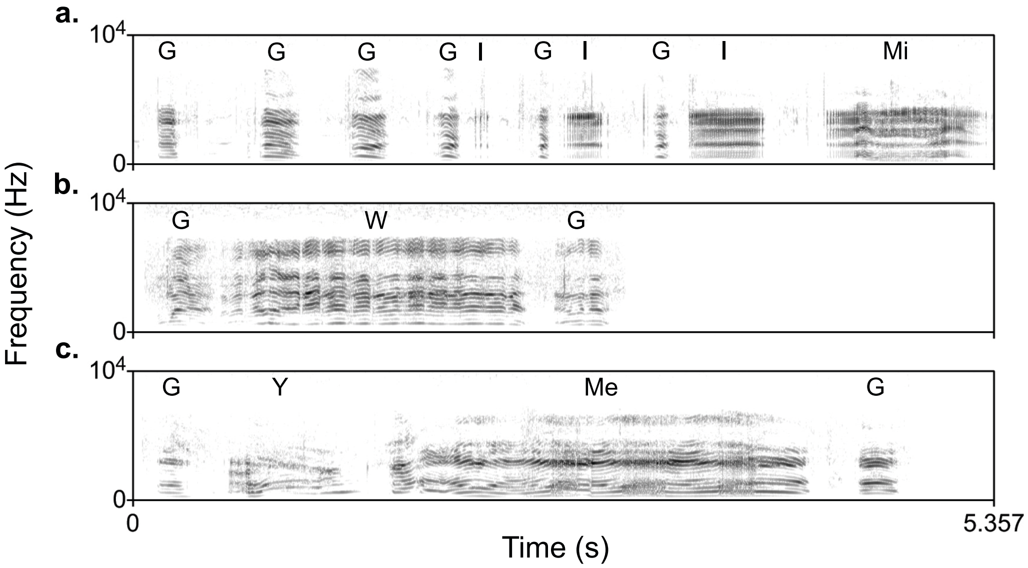


Figure 7.1. **Example spectrograms of vocal sequences produced during (a) matched-control sessions and (b-c) post-conflict sessions**. Uppercase letters above each utterance represent the call types: (G) exhaled grunt, (I) inhaled grunt, (Mi) inhaled moan, (W) wobble, (Y) vocal yawn, and (Me) exhaled moan. Spectrograms were made in Pratt (version 5.2.29).

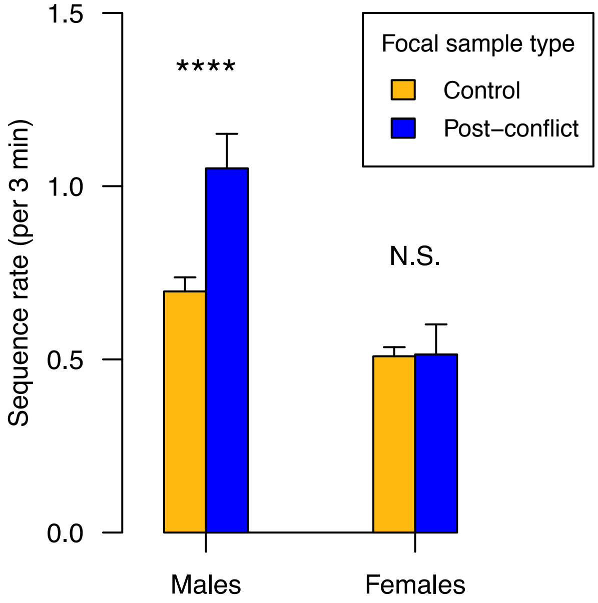


Figure 7.2. **Male and female vocal sequence rates in the 3 minutes after a 10-min 'conflict-free' control period or a conflict during focal samples**. Bars and whiskers represent means plus two SEM. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001, and p > 0.05 denoted as N.S.

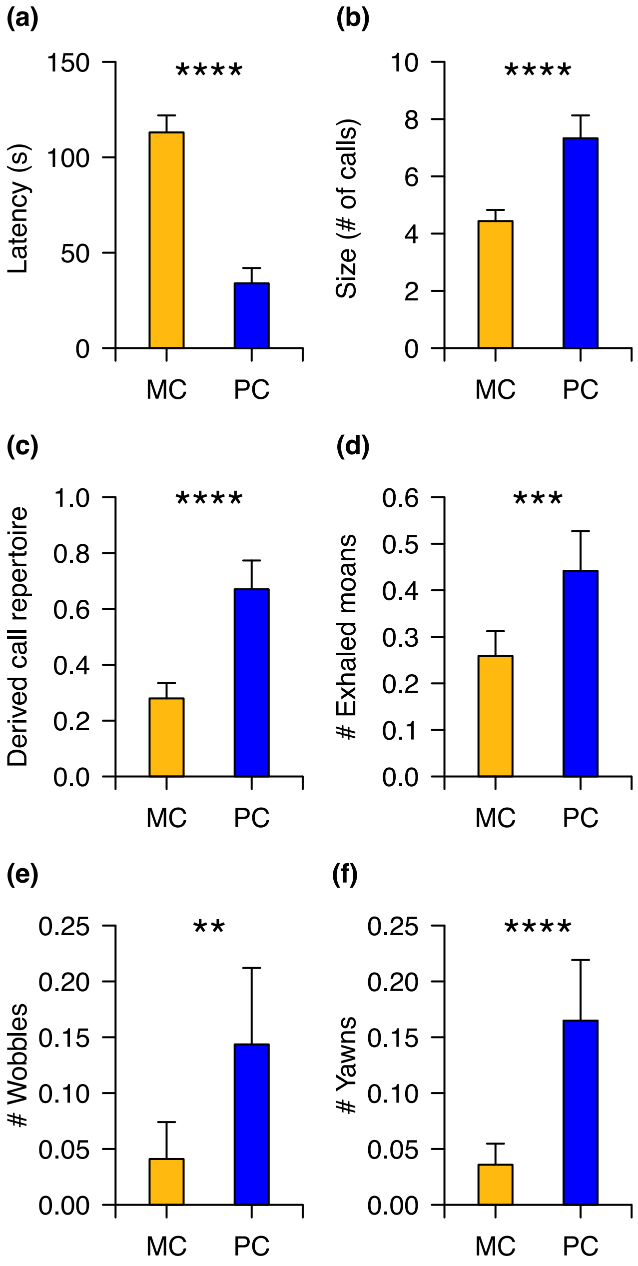


Figure 7.3. **Male vocal sequence production and complexity during matched-control (MC) and post-conflict (PC) video sessions**. Latency to give a sequence (a) was defined as the time after the start of the video (MC) or end of the conflict (PC). Bars and whiskers represent means plus two SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

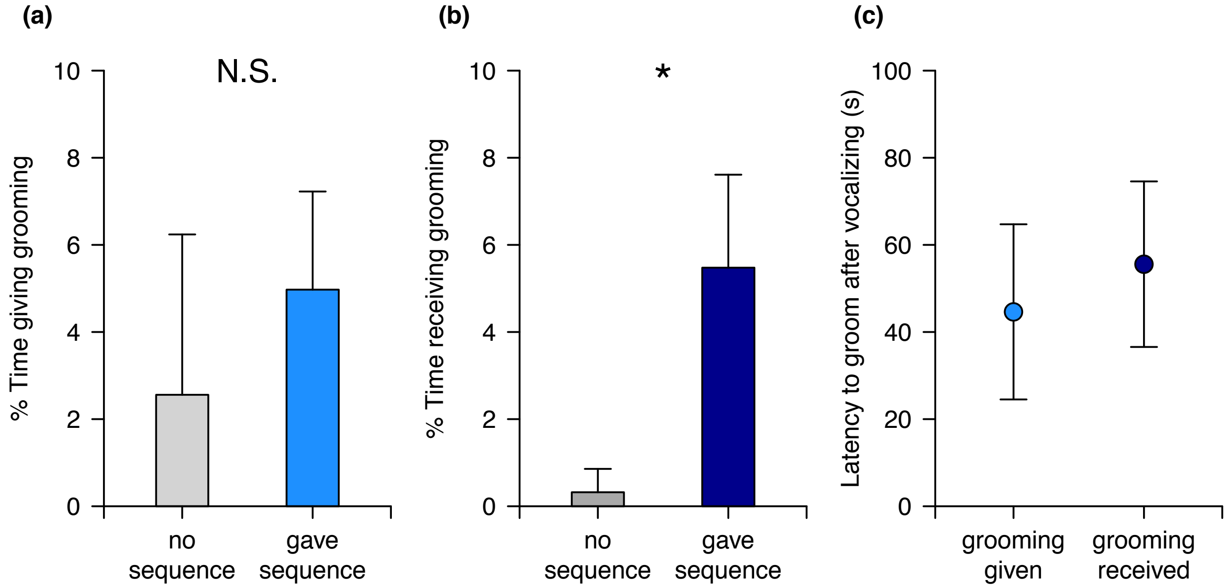


Figure 7.4. **Comparisons of post-conflict grooming behavior with unit females based on whether males produced a vocal sequence within the minute after conflicts**. Percent time (a) giving and (b) receiving grooming were calculated for the 3 min after the conflict. (c) Latencies to groom were calculated as the time between the first sequence and first grooming bout. Bars/points and whiskers represent means plus/minus two SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, and $p > 0.05$ denoted as N.S.

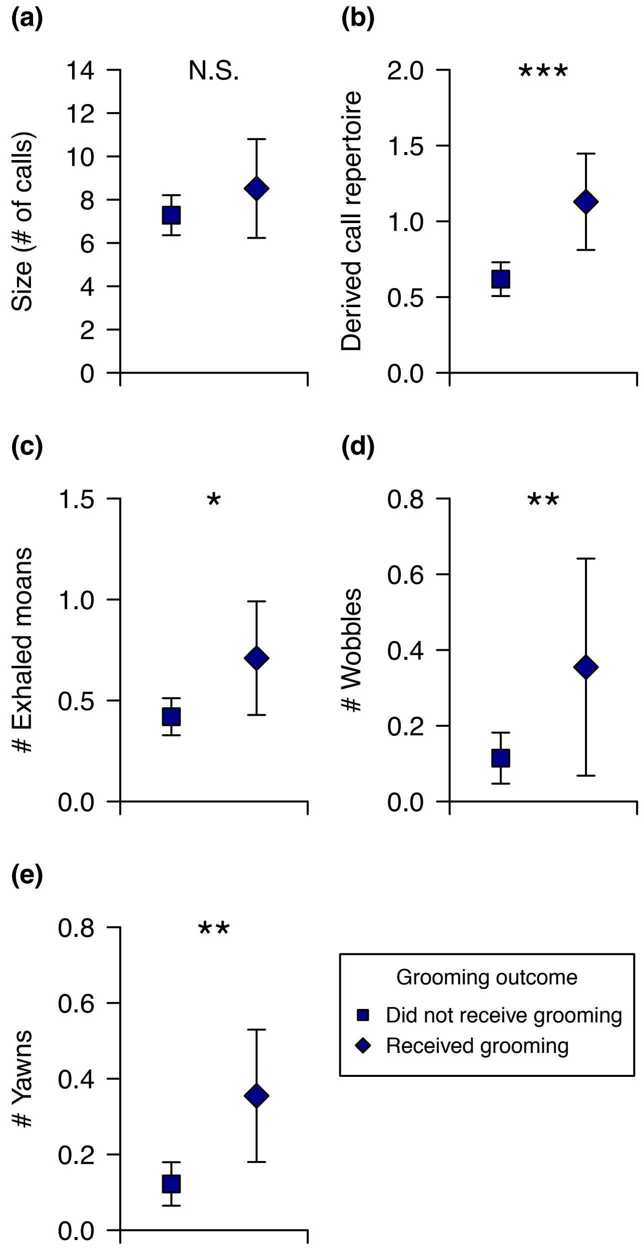


Figure 7.5. **Comparisons of grooming outcomes based on the (a) size and complexity features (b-e) of the first sequences produced within a minute after conflicts**. Points and whiskers represent means plus/minus two SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, and $p > 0.05$ denoted as N.S.

Chapter 8:

Vocal complexity varies with male quality in wild geladas

Abstract

The honest signaling hypothesis of sexual selection predicts that extravagant and complex vocal displays advertise caller quality. A wide range of literature on avian species supports this hypothesis, but we are lacking non-avian examples. Here, we investigate whether the complexity of gelada (*Theropithecus gelada*) male vocal sequences may function as a quality signal. Given that gelada males produce complex vocal sequences during affiliative interactions with females in their harem-like units, we tested the hypothesis that variation in sequence production (i.e., rate and sequence size) and complexity (e.g., number of derived call types per sequence) can signal a male's potential as a high-quality mate. Specifically, we compared sequence production and complexity in males thought to be higher and lower quality in gelada society (e.g., male status). We also examined correlations between sequence production and complexity and measures of fitness. We found that higher quality males (e.g., leader males) produce complex vocal sequences at higher rates. We did not find evidence that sequence production or complexity was related to fitness. Our findings are consistent with the hypothesis that gelada male vocal complexity is a quality signal, adding a new non-avian contribution to comparative study of extravagant displays in light of sexual selection.

Introduction

The honest signaling hypothesis of sexual selection states that extravagant and complex (i.e., multi-component) displays evolve because they are honest advertisements of individuals' quality or condition to potential mates and rivals (Johnstone, 1995; Zahavi, 1975, 1977). Vocalizations are one of the primary forms of displays in birds, amphibians and mammals. These vocal displays can function as quality signals if there are associated costs or constraints that prevent lower-quality individuals from making them (Johnstone, 1995; Zahavi, 1975). One of the best studied forms of extravagance in vocal displays is that of 'vocal complexity' or repertoire size in sequences of sounds (Byers & Kroodsma, 2009; Catchpole, 1987; Nowicki & Searcy, 2004, 2005; Searcy, 1992; Searcy, Nowicki, & Peters, 1999). A wide range of literature in songbirds (i.e., Passiformes) suggests that the ability to combine together a diverse repertoire of sounds characterizes individuals, typically males, who are high quality mates or competitors (Boogert, Giraldeau, & Lefebvre, 2008; Darolová, Krištofík, Hoi, & Wink, 2012; DeVoogd, 2004; Kagawa & Soma, 2013; Spencer, Buchanan, Goldsmith, & Catchpole, 2004).

While these associations between vocal complexity and male quality are quite common in songbirds, we are currently lacking comparable models in other lineages. Several non-avian species produce varied combinations of sounds in contexts relevant for inter- or intra-sexual selection: humans (Gao et al., 2017; Lange, Hennighausen, Brill, & Schwab, 2016; Lange, Zaretsky, Schwarz, & Euler, 2014; Rosenberg & Tunney, 2008), gibbons (Cowlishaw, 1992, 1996; Mitani, 1988), bats (Behr & von Helversen, 2004; Davidson & Wilkinson, 2004), mice (Chabout, Sarkar, Dunson, & Jarvis, 2015;

Holy & Guo, 2005), and frogs (Kime, Rand, Kapfer, & Ryan, 1998; Rand & Ryan, 1981; Ryan, 1983). However, direct associations between vocal complexity and caller quality are currently missing. Instead, acoustic properties such as frequency modulation, calling rate, call duration and amplitude are shown to act as quality signals in vocal displays (Barelli, Mundry, Heistermann, & Hammerschmidt, 2013; Benítez, le Roux, Fischer, Beehner, & Bergman, 2016; Charlton, Reby, & McComb, 2007; Fischer, Hammerschmidt, Cheney, & Seyfarth, 2002; Fischer, Kitchen, Seyfarth, & Cheney, 2004; Kitchen, Seyfarth, Fischer, & Cheney, 2003; Neumann, Assahad, Hammerschmidt, Perwitasari-Farajallah, & Engelhardt, 2010).

In this report, we seek to expand our understanding of vocal complexity as a quality signal by studying its variation in geladas (*Theropithecus gelada*), a non-human primate. There are two reasons why geladas make for an excellent study species for this question: their social system and their production of complex vocal displays. First, the gelada social system results in clear variation in male quality. Geladas live in large communities that are composed of smaller harem-like reproductive units with a dominant ‘leader’ male, up to a few subordinate ‘follower’ males and 2-11 related females and their immature offspring (Snyder-Mackler, Beehner, & Bergman, 2012). Unit males father all the offspring for the females in their units, and for units with follower males, the leaders father the majority (~85%) of the offspring (Snyder-Mackler, Alberts, & Bergman, 2012). Thus, there are clear ‘high quality’ (leaders) and ‘low quality’ (followers) males. Moreover, leader male quality is thought to deteriorate across tenure. Males are in their prime upon taking over units from the previous leader male, but their ability to maintain their position drops as tenure progresses (Snyder-Mackler, Alberts, et al., 2012). Thus, even within leaders, males earlier in their tenures are thought to be higher quality than males later in their tenures.

Second, gelada males are known for producing long and complex vocal sequences that are thought to function in facilitating social interactions with unit females (Aich, Moos-Heilen, & Zimmermann, 1990; Richman, 1987). ‘Complexity’ in gelada male vocal communication comes from their ability to produce a diversified vocal repertoire with a suite of five call types not found in their close baboon relatives (Gustison, le Roux, & Bergman, 2012). Three of the unique or ‘derived’ call types are thought to be the most salient and contain acoustic properties that make them difficult to produce (Gustison & Bergman, n.d.-a, 2016). These derived call types are combined with a shared call type (grunts) into sequences ranging anywhere from a single call to over twenty calls that contain none to all the derived call types (Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016). Vocal sequences containing more derived call types – specifically the three most salient derived calls – are produced during affiliative social interactions with unit members (especially females) and elicit grooming behavior from females (Gustison & Bergman, n.d.-b). In addition, females respond more strongly to playbacks of male vocal sequences that contain one of these three derived call types (Gustison & Bergman, 2016). Taken together, these data suggest that males who produce more complex vocal sequences (i.e., containing more derived call types) have stronger relationships with unit females.

Here, we investigate whether vocal sequence complexity maps on to male quality in male geladas. First, we test the prediction that leader males produce more complex vocal sequences than do followers. Second, we test the prediction that leader males early in their tenure produce more complex vocal sequences than leaders later in their tenure. In addition to investigating ‘complexity’ in gelada male vocal sequences, we also explore more generic measures of vocal production, specifically sequence rate and sequence size (number of calls in a sequence). Finally, we investigate whether sequence production or complexity at the start of leader tenures predicts the indirect and direct measures of fitness.

Methods

Study site and subjects

Data for this study come from adult geladas across three different bands in one wild community (about 1,200 individuals) living in the Sankaber area of the Simien Mountains National Park, Ethiopia (Snyder-Mackler, Beehner, et al., 2012). This population has been under intensive behavioral study since January 2006, and study subjects are fully habituated to human observers on foot (approach distances less than 3 m) (Bergman, 2010). The gelada research was approved by the University Committee on Use and Care of Animals at the University of Michigan and adhered to the laws and guidelines of Ethiopia.

Behavioral observations and acoustic sampling

From January 2009 to June 2016, we collected 977 hr of 15 minute focal samples from 85 adult male geladas (41 leaders, 24 followers, and 20 males who were leaders and followers during different parts of the study period). Focal samples included all-occurrence sampling (Altmann, 1974) of vocal, agonistic, and affiliative behaviors on a palm pilot. In addition, 2,128 vocal sequences (10,350 calls) were recorded from 49 males (30 leaders, 15 followers, and 4 males who were both leaders and followers) on an ad libitum basis from 2008-2009 and 2013-2014.

We define a vocal sequence as one or more discrete calls made by the same male and, for those sequences with two or more calls, separated by an inter-call interval of less than 2.983 s (Gustison & Bergman, n.d.-c). Avisoft SAS-Lab Pro (version 5.1.12, Avisoft Bioacoustics) was used to generate spectrograms with a Fast Fourier transformation (size of 1,024 points) and label the start times, end times, and call types of all calls in a vocal sequence. These call types included: exhaled grunts, inhaled grunts, exhaled moans, inhaled moans, wobbles and yawns and are described in previous studies (Gustison & Bergman, 2016; Gustison, le Roux, & Bergman, 2012; Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016). The size (total number of calls) and complexity of each sequence was determined following previous methods (Gustison & Bergman, n.d.-c). Complexity was measured by derived call repertoire size and the total number of each derived call type (exhaled moans, wobbles and yawns) per sequence.

Data analysis

We used Linear Mixed Models (LMM) to test for differences in sequence production and complexity based on male status (leader vs. follower) and the time since starting a leader tenure (number of days after takeover). In each LMM, sequence production (focal rate or sequence size) and complexity (derived call repertoire size per sequence or the number of exhaled moans, wobbles or yawns per sequence) was the dependent variable, and male status or time since beginning tenure was the fixed effect. In all LMMs, male identity, sampling date, and unit identity were random effects. LMMs were run using the function "lmer" of the R package "lme4" (Bates, Maechler, & Bolker, 2012). The "lmerTest" package was implemented to determine the significance of the LMM coefficients (Kuznetsova, Brockhoff, & Christensen, 2013).

Next, we tested for associations between leader male sequence production and complexity with a range of fitness parameters. Since leader male tenures range widely in length, we restricted sequence production and complexity measurements to the first six months of male tenures. In tests involving sequence rates we used males for which we had at least 4 focal samples, or 1 hour of observations, and we calculated the rate per 15 min. In tests involving sequence size and complexity, we used males for we had at least 4 sequence recordings, and we calculated the mean size and complexity measures. We investigated five fitness parameters: unit size at takeover (number of adult females), latency to first birth (for males that had at least one birth during tenure), tenure length (in years), female years (sum of the total number of years each female was in a male's unit), and total number of births. For birth-related fitness parameters, we only looked at births that occurred at least six months after the start of tenure, which is the gestation length of geladas (Roberts, Lu, Bergman, & Beehner, 2012), so that we could be confident that the offspring were not from a previous leader male. We could be confident that these offspring born after 6 mo were from the leader males, as extra-pair offspring in geladas are extremely rare (le Roux, Snyder-Mackler, Roberts, Beehner, & Bergman, 2013; Snyder-Mackler, Alberts, et al., 2012). First, we carried out Pearson correlations to test whether sequence production or complexity measures were associated with the unit size at takeover (number of adult females) or latency to the first birth. Then, we carried out Pearson correlations to test whether sequence production or complexity measures were related tenure length, female years or total number of births. We constrained these latter correlations to males for which we had data on the entire tenure (i.e., from takeover to takeover).

All calculations and statistical tests were carried out in R 3.3.0 (R Development Core Team, 2016). Critical values were set at alpha = 0.05, and all tests were two-tailed.

Results

Comparison of leader and follower vocal sequence performance

We used a series of LMMs to compare leader and follower males in their sequence rate, size (number of calls per sequence) and complexity (derived call repertoire size and number of each derived call type per sequence). We found that leaders produce sequences at higher rates (per 15 min) than followers (Estimate ± SE = 0.518 ± 0.180, $t = 2.886$, $p = 0.0040$; Figure 8.1a), and leaders' sequences were larger in size (Estimate ± SE = 1.678 ± 0.405, $t = 4.143$, $p < 0.0001$; Figure 8.1b) and contained more derived call types (Estimate ± SE = 0.276 ± 0.057, $t = 4.845$, $p < 0.0001$; Figure 8.1c) than follower sequences. Specifically, leaders' sequences contained more exhaled moans (Estimate ± SE = 0.225 ± 0.050, $t = 4.481$, $p < 0.0001$; Figure 8.1d) and vocalized yawns (Estimate ± SE = 0.088 ± 0.027, $t = 3.257$, $p = 0.0021$; Figure 8.1f). The only sequence feature for which we did not detect a difference was the number of wobbles per sequence in leaders compared to followers (Estimate ± SE = -0.033 ± 0.052, *t* = -0.642, *p* = 0.5216; Figure 8.1e).

Changes in vocal sequence performance across leader tenure

Leader tenures ranged from a couple weeks to 8 years. We used a series of LMMs to test whether there were any longitudinal changes in vocal sequence production and complexity across leader tenures. Sequence rate (per 15 min focal) was negatively associated with the number of days after tenure began (Intercept = 4.404 ± 0.251, Estimate ± SE = $-5.057 \times 10^{-4} \pm 1.654 \times 10^{-4}$, *t* = -3.058, *p* = 0.0023; Figure 8.2a). However, we found no evidence that sequence size was associated with the number of days after tenure began (Intercept = 5.405 ± 0.309, Estimate ± SE = $-1.081 \times 10^{-4} \pm 3.290 \times 10^{-4}$, *t* = -0.328, *p* = 0.7429; Figure 8.2b). On the other hand, the derived call repertoire size per sequence was negatively associated with the number of days after tenure began (Intercept = 0.625 ± 0.044, Estimate ± SE = $-1.419 \times 10^{-4} \pm 0.494 \times 10^{-4}$, *t* = -2.871, *p* = 0.0052; Figure 8.2c). Similarly, the number of wobbles per sequence was negatively associated with the number of days after tenure began (Intercept = 0.109 ± 0.018, Estimate ± SE = $-6.519 \times 10^{-5} \pm 2.120 \times 10^{-5}$, *t* = -3.075, *p* = 0.0034). We found no clear evidence that the number of exhaled moans (Intercept = 0.436 ± 0.038, Estimate ± SE = $-6.744 \times 10^{-5} \pm 4.172 \times 10^{-5}$, *t* = -1.616, *p* = 0.1086) or number of yawns (Intercept = 0.163 ± 0.022, Estimate ± SE = $-4.054 \times 10^{-5} \pm 2.483 \times 10^{-5}$, *t* = -1.633, *p* = 0.1086) were associated with the number of days after tenure began.

Correlations between vocal sequence performance and fitness parameters

We found no evidence that sequence production or complexity was related to fitness parameters (Table 8.1). Although none of the correlations were statistically significant, the largest effect sizes were between sequence measures and latency to first birth. Sequence rate had a negative r value suggesting that males that produced sequences at higher rates tended to have births faster (Figure 8.3a). In contrast, sequence size and derived call repertoire had positive r values in that males that produced longer sequences containing more derived call types tended to have births later (Figure 8.3b-c).

Discussion

To our knowledge, this study is the first to systematically demonstrate that ‘vocal complexity’ can signal male quality in a non-avian species. Specifically, dominant leader male geladas produce vocal sequences that are more complex (i.e., contain more derived call types) than subordinate follower males, and vocal complexity decreases throughout leaders’ tenure. Even other measures of vocal performance (sequence rate and sequence size) differed based on male quality. Compared to followers, leaders vocalized at higher rates and used longer sequences compared to followers, and sequence rate dropped as leaders’ tenures progressed. Taken together, these findings support the honest signaling hypothesis predicting that extravagant, multi-component displays advertise quality in contexts relevant for sexual selection (Johnstone, 1995; Zahavi, 1975, 1977). Moreover, these data align with an extensive body of research on courtship song in Passerine birds (Catchpole, 1987; Nowicki, Searcy, Hughes, & Podos, 2001; Searcy, 1992) and extends the work on the exaggerated acoustic properties of male-male competitive displays in non-human primates (reviewed in: Delgado, 2006).

Several non-avian taxa produce varied combinations of vocalizations in contexts relevant for sexual selection, but we are unaware of studies showing a direct link with caller quality (Behr & von Helversen, 2004; Chabout et al., 2015; Cowlishaw, 1996; Kime et al., 1998; Rosenberg & Tunney, 2008). Although studies on non-avian taxa sometimes show that females prefer males who produce complex vocal sequences (e.g., Chabout et al., 2015; Kime et al., 1998), female preference for complex vocal displays does not guarantee that the callers also differ in quality (Collins, 1999; Endler & Basolo, 1998; Ryan, 1998). For example, exploitation of sensory bias is thought to underlie female preference of complex calls in Túngara frogs (*Physalaemus pustulosus*), and males that differ in quality do not differ in complex call production (Bernal, Page, Rand, & Ryan, 2007). Prior research on gelada female responses to complex male vocal sequences showed rudimentary support for either explanation. Female geladas looked longer at playbacks of male sequences containing one of the three most salient derived call types than playbacks of grunt-only sequences, but it was unclear if this was because callers of complex sequences were perceived as higher quality, signaled an affiliative social interaction, or because the derived calls were more acoustically salient that grunts (Gustison & Bergman, 2016). The present study provides support for the first two explanations, that females may be responding to signals of male quality and context.

For complex vocal signals to act as honest advertisements of quality, they must be difficult to produce. The current study, combined with previous work, supports this prediction in geladas. The three most salient derived call types contain more potentially costly acoustic properties like prolonged duration, coordination of sophisticated facial movements, and modulation of fundamental and formant frequency profiles (Gustison & Bergman, n.d.-a, 2016). Even long sequences containing several calls are thought to be more difficult to produce, which results in compression as a possible coping mechanism (Gustison et al., 2016). The current study supports the prediction that these complex vocal sequences are difficult to produce given that we saw robust differences across males differing in quality. This is the second study to demonstrate that wild gelada vocal displays may act as male quality signals. The first study focused on bouts of louds calls given by males during male-male competition, finding that measures of male quality (e.g., leader vs. follower status and estimated age) predicted acoustic properties of display calls (Benítez et al., 2016). Higher quality males produced bouts of display calls that were longer, lower in overall frequency, and covered a larger frequency range. Taken together, there is now good evidence that gelada male vocal displays may act as quality signals in both inter- and intra-sexual selection contexts.

Our findings also have implications for understanding the social functions of gelada male vocal sequences. Previous work indicated that gelada males produce complex sequences to facilitate affiliative interactions with females during periods of stability (Gustison & Bergman, n.d.-c) and instability (Gustison & Bergman, n.d.-b). Moreover, production of complex sequences appeared to elicit female investment in the male-female social bond. One of the outstanding questions was why males, but not females, produce these complex vocal sequences (Gustison & Bergman, n.d.-c). One hypothesis was that gelada males are in the unique position of having to build new social relationships upon taking over a reproductive unit. In support of this hypothesis, our findings show that leaders produce a higher rate of sequences containing more derived call types earlier in tenure. This finding also reveals that the usage of gelada complex sequences is more analogous to avian and human communication than previously thought. Like geladas, male zebra finches (*Taeniopygia guttata*) engage in complex songs when forming a long-term pair bond with a female, but then they transition to using simpler calls for the duration of the partnership (Elie et al., 2010; Tomaszycki & Adkins-Regan, 2005, 2006). Communication changes also happen in a similar direction during the development of intimate human relationships (Punyanunt-Carter, 2004).

It is still unknown whether variation in gelada leader male production and complexity of sequences aligns with indirect or direct fitness benefits. We found no evidence that a leader male's vocal performance during the beginning of tenure was correlated with unit size, tenure length, or the latency to and number of births. Two potential explanations may explain this lack of association. On the one hand, it is possible that we did not have a large enough sample size to detect an association. There were a limited number of males for which we could collect substantial behavioral observations and acoustic recordings early in tenure and continue following them until their tenure ended. On the other hand, it may also be the case that there is little variability in the fitness potential of new leaders because males must successfully compete to become leaders, and therefore, may already be high quality. Before untangling these possible explanations, intense longitudinal studies of complex vocal behaviors throughout leader male tenures will be needed.

Acknowledgements

We thank the Ethiopian Wildlife Conservation Authority (EWCA) as well as the wardens and staff of the Simien Mountain National Park for permission and support in conduction research on geladas. We also thank all the members of the University of Michigan Gelada Research Project (UMGRP) for assistance in data collection and valuable insight on analyses and interpretation. MLG was supported by grants from the National Geographic Society (9122-12, W304-14), Leakey Foundation, University of Michigan, and Petridish (primary supporter: John Allen). UMGRP was supported by the Wildlife Conservation Society (SSF 67250), the National Geographic Society (8100-06), the Leakey Foundation, the National Science Foundation (BCS-0715179, BCS-0962118, IOS- 1255974), and the University of Michigan.

References

Aich, H., Moos-Heilen, R., & Zimmermann, E. (1990). Vocalizations of adult gelada baboons (*Theropithecus gelada*): Acoustic structure and behavioural context. *Folia Primatologica*, *55*, 109–132.

Altmann, J. (1974). Observational study of behavior : Sampling methods. *Behaviour*, *49*, 227–267.

Barelli, C., Mundry, R., Heistermann, M., & Hammerschmidt, K. (2013). Cues to androgens and quality in male gibbon songs. *PloS One*, *8*, e82748.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. Retrieved from http://cran.r-project.org/ package=lme4

Behr, O., & von Helversen, O. (2004). Bat serenades—complex courtship songs of the sac-winged bat (*Saccopteryx bilineata*). *Behavioral Ecology and Sociobiology*, *56*, 106–115.

Benítez, M. E., le Roux, A., Fischer, J., Beehner, J. C., & Bergman, T. J. (2016). Acoustic and temporal variation in gelada (Theropithecus gelada) loud calls advertise male quality. *International Journal of Primatology*. http://doi.org/10.1007/s10764-016-9922-0

Bergman, T. J. (2010). Experimental evidence for limited vocal recognition in a wild primate: implications for the social complexity hypothesis. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 3045–53.

Bernal, X. E., Page, R. A., Rand, A. S., & Ryan, M. J. (2007). Natural history miscellany - Cues for eavesdroppers: Do frog calls indicate prey density and quality? *The American Naturalist*, *169*, 409–415.

Boogert, N. J., Giraldeau, L.-A., & Lefebvre, L. (2008). Song complexity correlates with learning ability in zebra finch males. *Animal Behaviour*, *76*, 1735–1741.

Byers, B. E., & Kroodsma, D. E. (2009). Female mate choice and songbird song repertoires. *Animal Behaviour*, *77*, 13–22.

Catchpole, C. K. (1987). Bird song, sexual selection and female choice. *Trends in Ecology and Evolution*, *2*, 94–97.

Chabout, J., Sarkar, A., Dunson, D. B., & Jarvis, E. D. (2015). Male mice song syntax depends on social contexts and influences female preferences. *Frontiers in*

Behavioral Neuroscience, *9*, 1–16.

Charlton, B. D., Reby, D., & McComb, K. E. (2007). Female red deer prefer the roars of larger males. *Biology Letters*, *3*, 382–385.

Collins, S. A. (1999). Is female preference for male repertoires due to sensory bias? *Proceedings of the Royal Society B: Biological Sciences*, *266*, 2309–2314.

Cowlishaw, G. (1992). Song Function in Gibbons. *Behaviour*, *121*, 131–153.

Cowlishaw, G. (1996). Sexual selection and information content in gibbon song bouts. *Ethology*, *102*, 272–284.

Darolová, A., Krištofík, J., Hoi, H., & Wink, M. (2012). Song complexity in male marsh warblers: does it reflect male quality? *Journal of Ornithology*, *153*, 431–439.

Davidson, S. M., & Wilkinson, G. S. (2004). Function of male song in the greater white-lined bat, *Saccopteryx bilineata*. *Animal Behaviour*, *67*, 883–891.

Delgado, R. a. (2006). Sexual Selection in the loud calls of male primates: Signal content and function. *International Journal of Primatology*, *27*, 5–25.

DeVoogd, T. J. (2004). Neural constraints on the complexity of avian song. *Brain, Behavior and Evolution*, *63*, 221–32.

Elie, J. E., Mariette, M. M., Soula, H. A., Griffith, S. C., Mathevon, N., & Vignal, C. (2010). Vocal communication at the nest between mates in wild zebra finches: a private vocal duet? *Animal Behaviour*, *80*, 597–605.

Endler, J. A., & Basolo, A. L. (1998). Sensory ecology, receiver biases and sexual selection. *Trends in Ecology and Evolution*, *13*, 415–420.

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

Fischer, J., Kitchen, D. M., Seyfarth, R. M., & 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*, 140–148.

Gao, Z., Gao, S., Xu, L., Zheng, X., Ma, X., Luo, L., & Kendrick, K. M. (2017). Women prefer men who use metaphorical language when paying compliments in a romantic context. *Scientific Reports*, *7*, 40871.

Gustison, M. L., & Bergman, T. J. (In Review: n.d.-a). Divergent acoustic properties of gelada and baboon vocalizations and their implications for the evolution of human speech. *Journal of Language Evolution*.

Gustison, M. L., & Bergman, T. J. (n.d.-b). Gelada males use complex vocal sequences to manage conflicts.

Gustison, M. L., & Bergman, T. J. (n.d.-c). The social functions of complex call sequences in wild geladas.

Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Scientific Reports*, *6*, 19680.

Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (Theropithecus gelada) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1847–59.

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*, E2750–E2758.

Holy, T. E., & Guo, Z. (2005). Ultrasonic songs of male mice. *PLoS Biology*, *3*, e386.

Johnstone, R. A. (1995). Honest advertisement of multiple qualities using multiple signals. *Journal of Theoretical Biology*, *177*, 87–94.

Kagawa, H., & Soma, M. (2013). Song performance and elaboration as potential indicators of male quality in Java sparrows. *Behavioural Processes*, *99*, 138–144.

Kime, N. M., Rand, A. S., Kapfer, M., & Ryan, M. J. (1998). Consistency of female choice in the túngara frog: a permissive preference for complex characters. *Animal Behaviour*, *55*, 641–649.

Kitchen, D. M., Seyfarth, R. M., Fischer, J., & Cheney, D. L. (2003). Loud calls as indicators of dominance in male baboons (*Papio cynocephalus ursinus*). *Behavioral Ecology and Sociobiology*, *53*, 374–384.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). Retrieved from http://cran.r-project.org/package=lmerTest

Lange, B. P., Hennighausen, C., Brill, M., & Schwab, F. (2016). Only Cheap Talk after All? New Experimental Psychological Findings on the Role of Verbal Proficiency in

Mate Choice. *Psychology of Language and Communication*, *20*, 1–22.

Lange, B. P., Zaretsky, E., Schwarz, S., & Euler, H. A. (2014). Words Won't Fail. *Journal of Language and Social Psychology*, *33*, 482–499.

le Roux, A., Snyder-Mackler, N., Roberts, E. K., Beehner, J. C., & Bergman, T. J. (2013). Evidence for tactical concealment in a wild primate. *Nature Communications*, *4*, 1462.

Mitani, J. C. (1988). Male gibbon (Hylobates agilis) singing behavior: Natural history, song variations and function. *Ethology*, *79*, 177–194.

Neumann, C., Assahad, G., Hammerschmidt, K., Perwitasari-Farajallah, D., & Engelhardt, A. (2010). Loud calls in male crested macaques, Macaca nigra: a signal of dominance in a tolerant species. *Animal Behaviour*, *79*, 187–193.

Nowicki, S., & Searcy, W. A. (2004). Song function and the evolution of female preferences: why birds sing, why brains matter. *Annals of the New York Academy of Sciences*, *1016*, 704–23.

Nowicki, S., & Searcy, W. A. (2005). Song and mate choice in birds: How the development of behavior helps us understand function. *The Auk*, *122*, 1–14.

Nowicki, S., Searcy, W. A., Hughes, M., & Podos, J. (2001). The evolution of bird song: male and female response to song innovation in swamp sparrows. *Animal Behaviour*, *62*, 1189–1195.

Punyanunt-Carter, N. M. (2004). Satisfaction in marital and dating relationships. *Psychological Reports*, *95*, 1154–1160.

R Development Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Rand, A. S., & Ryan, M. J. (1981). The adaptive significance of a complex vocal repertoire in a neotropical frog. *Zeitschrift Für Tierpsychologie*, *57*, 209–214.

Richman, B. (1987). Rhythm and melody in gelada vocal exchanges. *Primates*, *28*(2), 199–223.

Roberts, E. K., Lu, A., Bergman, T. J., & Beehner, J. C. (2012). A Bruce effect in wild geladas. *Science*, *335*, 1222–1225.

Rosenberg, J., & Tunney, R. J. (2008). Human vocabulary use as display. *Evolutionary*

Psychology, *6*, 538–549.

Ryan, M. J. (1983). Sexual selection and communication in a Neotropical frog, *Physalaemus pustulosus*. *Evolution*, 37, 261-272.

Ryan, M. J. (1998). Sexual selection, receiver biases, and the evolution of sex differences. *Science*, *281*, 1999–2003.

Searcy, W. A. (1992). Song reportoire and mate choice in birds. *American Zoologist*, *32*, 71–80.

Searcy, W. A., Nowicki, S., & Peters, S. (1999). Song types as fundamental units in vocal repertoires. *Animal Behaviour*, *58*, 37–44.

Snyder-Mackler, N., Alberts, S. C., & Bergman, T. J. (2012). Concessions of an alpha male? Cooperative defence and shared reproduction in multi-male primate groups. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 3788–3795.

Snyder-Mackler, N., Beehner, J. C., & Bergman, T. J. (2012). Defining higher levels in the multilevel societies of geladas (*Theropithecus gelada*). *International Journal of Primatology*, *33*, 1054–1068.

Spencer, K. a, Buchanan, K. L., Goldsmith, A. R., & Catchpole, C. K. (2004). Developmental stress, social rank and song complexity in the European starling (*Sturnus vulgaris*). *Biology Letters*, *271 Suppl*, S121-3.

Tomaszycki, M. L., & Adkins-Regan, E. (2005). Experimental alteration of male song quality and output affects female mate choice and pair bond formation in zebra finches. *Animal Behaviour*, *70*, 785–794.

Tomaszycki, M. L., & Adkins-Regan, E. (2006). Is male song quality important in maintaining pair bonds? *Behaviour*, *143*, 549–567.

Zahavi, A. (1975). Mate selection-A selection for a handicap. *Journal of Theoretical Biology*, *53*, 205–214.

Zahavi, A. (1977). The cost of honesty. Further Remarks on the Handicap Principle. *Journal of Theoretical Biology*, *67*, 603–605.

Table 8.1. Results of Pearson correlation tests between leader male sequence production and complexity measures (from the first 6 months of tenure) and fitness measures.

Sequence measure	***Unit size at takeover***		***Latency to 1st birth (days)***		***Tenure length (years)***		***Female years***		***Total # of births***	
	r (df)	*p*	*r* (df)	*p*	*r* (df)	*p*	*r* (df)	*p*	*r* (df)	*p*
Rate per 15 min	0.12 (35)	0.4868	-0.20 (26)	0.3093	0.02 (21)	0.9462	-0.05 (21)	0.8300	0.06 (21)	0.7708
Mean size	-0.07 (12)	0.8072	0.33 (9)	0.3166	0.04 (7)	0.9230	-0.11 (7)	0.7821	-0.20 (7)	0.6037
Mean derived call repertoire	0.29 (12)	0.3211	0.22 (9)	0.5065	0.06 (7)	0.8739	0.03 (7)	0.9463	-0.12 (7)	0.7650
Mean # exhaled moans	0.32 (12)	0.2596	0.13 (9)	0.6965	-0.14 (7)	0.7254	-0.01 (7)	0.9838	0.01 (7)	0.9876
Mean # wobbles	0.10 (12)	0.7453	-0.08 (9)	0.8150	-0.42 (7)	0.2608	-0.46 (7)	0.2129	-0.42 (7)	0.2637
Mean # yawns	0.12 (12)	0.6719	0.25 (9)	0.4622	0.32 (7)	0.4068	0.21 (7)	0.5805	0.09 (7)	0.8092

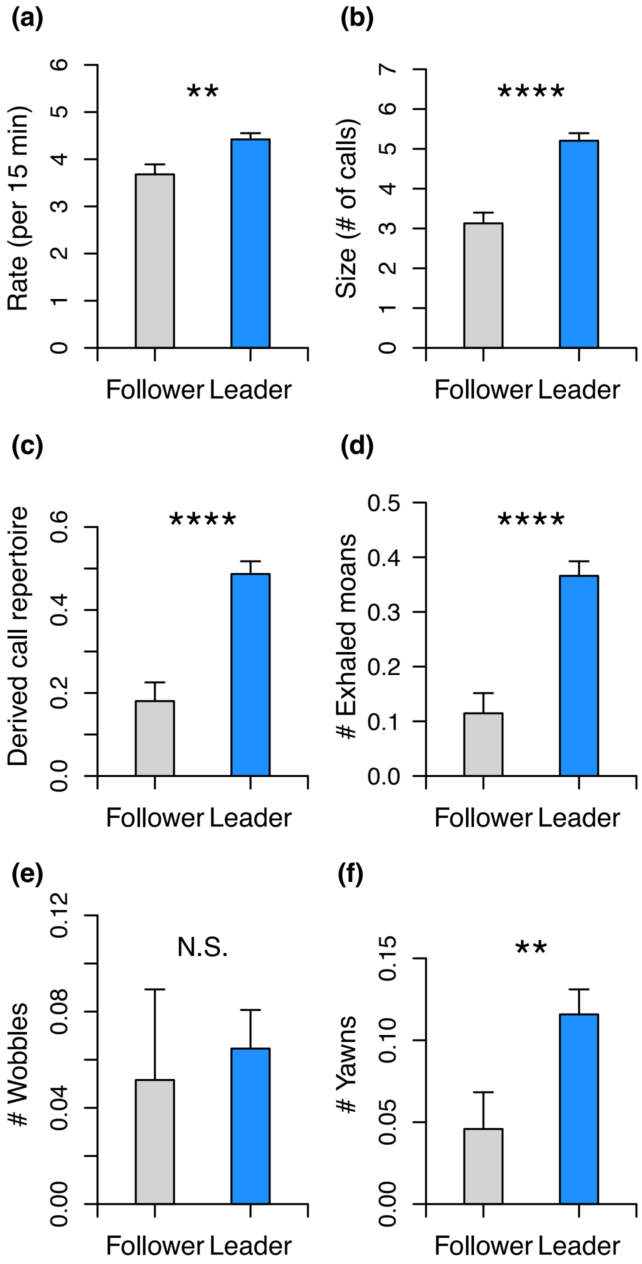


Figure 8.1. Sequence (a-b) production and (c-f) complexity in follower and leader males. Bars and whiskers represent means + 2 SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ and $p > 0.05$ denoted as N.S.

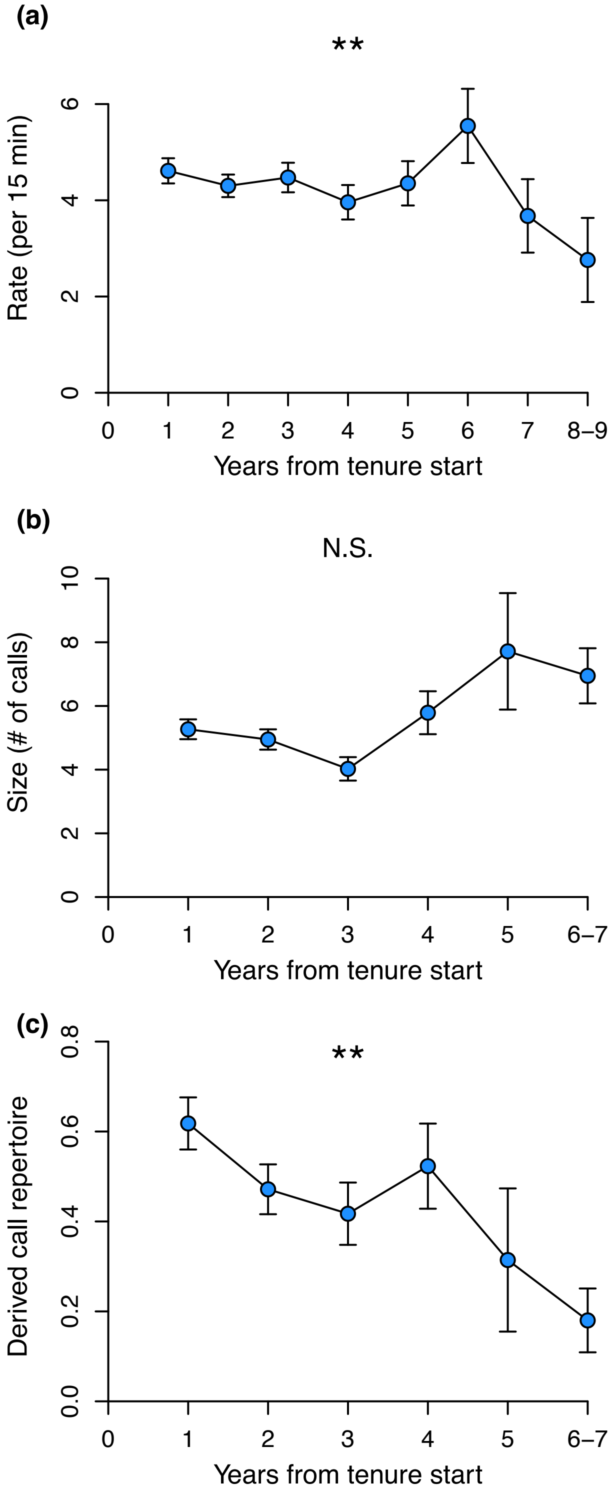


Figure 8.2. Sequence (a-b) production and (c) complexity throughout leader male tenure. Points and whiskers represent means + 2 SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ and $p > 0.05$ denoted as N.S.

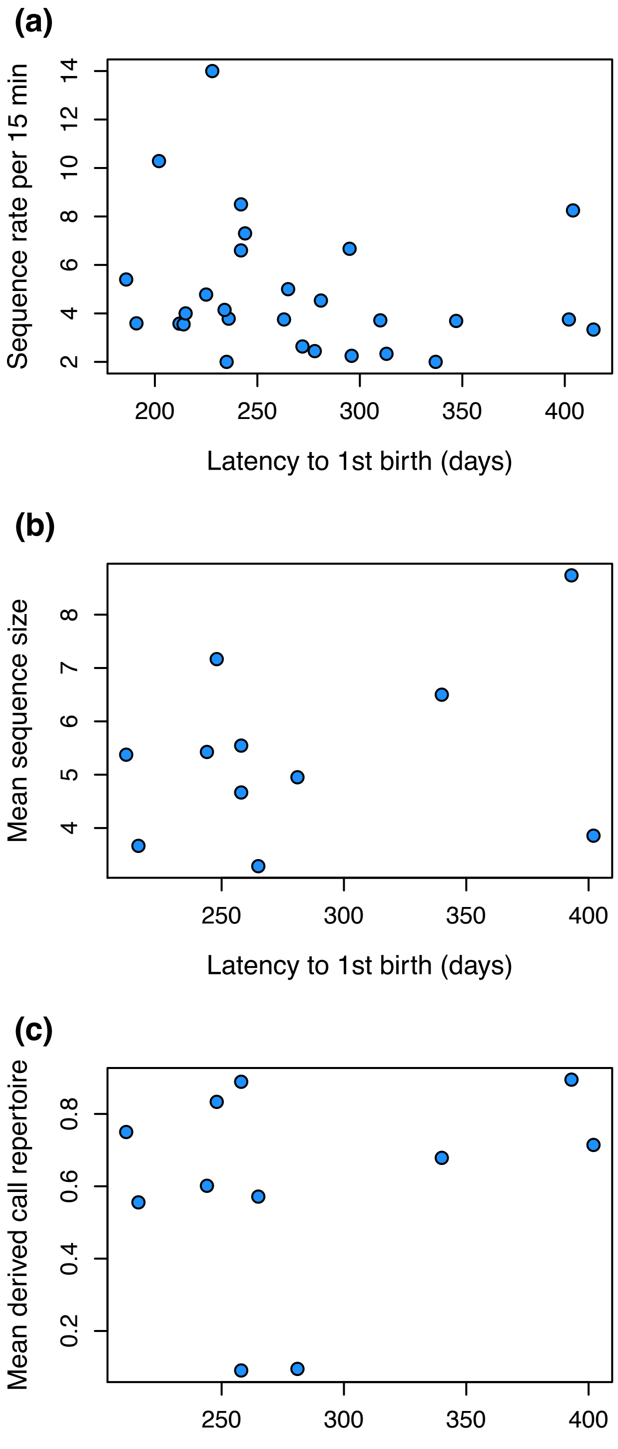


Figure 8.3. Associations between latency to 1st birth and sequence (a-b) production and (c) complexity in the first 6 months of leader male tenures. Each point represents a single tenure.

Chapter 9:

Conclusions

This dissertation has examined the phylogeny (Chapters 2-4) and function (Chapters 5-8) of vocal complexity in wild geladas. In doing so, this work provides a foundation for thinking about the ultimate origins of ‘complexity’ in vocal communication systems. Additionally, this work has provided key insights in three areas.

First, comparative work on the parallels between social complexity and vocal complexity have left us with the puzzle of how these two entities are linked. Our work provides the first thorough demonstration of how this can happen, and it sets up guidelines for how this link can be understood in other taxa. As a first step, it is critical to identify the specific features of vocal communication that are thought of as ‘complex’, and this identification can be helped by making comparisons between closely related species. In the case of geladas, comparing them to baboons helped us to determine which call types in their repertoire were unique and therefore created a larger vocal repertoire size (Gustison, le Roux, & Bergman, 2012). Then, we could focus on determining the social function of these unique ‘derived’ calls (Gustison & Bergman, 2016; Gustison et al., 2012). There are many ways that a complex communication system could function in a species thought to be socially complex. If a species is thought to be socially complex because of their large group sizes, then it is necessary to determine if complex vocalizations are used to combat conspecific noise, facilitate contact between individuals who are lost in the crowd, enhance the ability to compete for resources, and/or deal with an influx of potential predators, for example. In geladas, we found support for the idea that complex sequences help gelada males overcome the noise created by living in a large and dense community (Chapter 6). If a species is thought to be socially complex due to being incredibly gregarious, then it is necessary to determine if complex vocalizations function by replacing or facilitating specific types of affiliative interactions. In geladas, we found support for the idea that complex sequences facilitate gelada males' ability to engage in affiliative interactions with females, especially during periods of uncertainty (Chapters 6 and 7). These links between social complexity and vocal complexity in geladas may apply to other species, but this is not necessarily the case. To better understanding how social complexity drives vocal complexity (or vice versa) we need to know how this link manifests in a wide range of species.

Second, our work provides novel comparative data with which to make inferences about the origins spoken human language, the vocal communication system presumed to be the most complex (Fitch, 2000; Hauser et al., 2014). 'Complexity' in human language comes in many forms. The form that is potentially the easiest to compare to our non-human relatives is our phonological capacity to produce an impressive repertoire of sounds. Traditionally, our capacity to produce a wide range of sounds was thought to be a unique outcome of advanced anatomy among the primate lineage (Lieberman, Klatt, & Wilson, 1969). However, this view is being challenged by recent studies of vocal tract morphology in baboons and macaques (Boë et al., 2017; Fitch, de Boer, Mathur, & Ghazanfar, 2016). Our work fits into this larger discussion by showing that geladas can produce sounds that are even more modulated than their Papionini relatives – both in spectrum-based and temporal-based acoustic properties (Gustison & Bergman, n.d.). Our work, combined with recent studies on other Papionini species, suggests that the human phonological capacity is better explained by taxonomic differences in neural circuitry rather than vocal tract morphology.

Another potential form of complexity in spoken human language is our emergent capacity for 'discrete infinity', or the ability to produce an infinite range of messages from a finite repertoire of sounds (Hauser et al., 2014; Hauser, Chomsky, & Fitch, 2002). Our ability to do this comes from our application of syntactical rules to decode the organization of speech and use semantic rules to infer meaning. One of the major questions for understanding language origins is which came first: the ability to produce a structured string of sounds or the ability to map meaning onto specific sounds. A wealth of literature explores the syntax-like and semantic-like capacity of non-human vocal systems with the goal of trying to figure out the evolutionary history of these abilities (Berwick, Okanoya, Beckers, & Bolhuis, 2011; Collier, Bickel, van Schaik, Manser, & Townsend, 2014; Manser, 2013; Seyfarth et al., 2010). These non-human studies typically fall under one of two categories: those who focus on syntax-like ability in the absence of 'meaning' (e.g., the courtship songs of birds) and those who focus on semantic-like ability in both the presence and absence of proto-syntax (e.g., predator and food calls of non-human primates). Our work is best aligned with the idea that complex forms of vocal communication can evolve in the absence of intricate semantic content. This is comparable to the 'redundancy' property of human language – despite having over 20,000 word families, we only require 6,000-7,000 (Rosenberg & Tunney, 2008). Our work shows that geladas are a useful non-human primate model with which to investigate how 'extravagance' emerges (Chapter 8).

Another potential form of complexity in spoken human language is our capacity for pragmatics – often described as the ability to use context to modify the meaning of communication signals (Fitch, 2010). Our work shows that geladas may be a useful comparative model for understanding the role of pragmatics in complex vocal communication. Overall, the information content or 'meaning' of derived call types used by males during vocal sequences seemed to be relatively redundant – all derived call types appeared to act as signals of benign intent (Chapters 5-7). That said, there were subtle differences in their use that varied based on the social context. For example, vocal sequences containing more wobbles were predictive of male callers grooming female recipients, while vocal sequences containing more vocal yawns were predictive of female recipients grooming male callers. Meanwhile, the number of exhaled moans per sequence did not predict grooming outcomes (Chapter 6). On the other hand, vocal sequences containing more exhaled moans, wobbles or yawns, were predictive of female receipts grooming male callers (Chapter 7). These findings suggest that social context (fight or no fight) could potentially modify the meaning of complex vocal sequences in geladas.

Third, our work offers novel insight into the evolution of complexity in biological systems. One way it does this is by showing that the use of 'diversity' as an index of complexity has ecological validity. Complexity is a tricky property to quantify, and to cope with this limitation, scientists often use diversity as an indirect measure of complexity (D'Arnoldi, Foulley, & Ollivier, 1998; Kershenbaum, 2014; Page, 2010;

Shannon, 1948; Weitzman, 1992). Diversity can be measured on a variety of levels in multi-component systems – from the number of unique attributes and entropy (the studies in this dissertation) to the degree of distance between components. Our work suggests that relatively simple measures of diversity can be used to better understand the origins of 'complexity' in communication systems.

It is important to note that there are trade-offs to producing complex forms of communication. It is not a given that complex communications are 'better'. For example, signals that rely on brevity can be more efficient in transmitting a message than signals that are more complex (Kempe, Puts, & Cárdenas, 2013; Luo et al., 2013; Semple, Hsu, & Agoramoorthy, 2010). Complex signals may also require more physiological and developmental investment, which be more costly than beneficial (Suzuki, Ikebuchi, Bischof, & Okanoya, 2014). Our work suggests that the social benefits of vocal complexity can outweigh potential costs, at least in certain contexts and environments.

A final way our work offers insight into the evolution of complexity is by supporting the idea that universal principles may dictate how it emerges in biological systems (Gustison & Bergman, n.d.; Gustison, Semple, Ferrer-i-Cancho, & Bergman, 2016; Poincaré, 1913; Popper, 1934). Menzerath's law, for example, is a principle thought to underpin the organizational structure of many combinatorial systems. It simply states that "the greater the whole, the smaller its constituents" (Altmann, 1980; Köhler, 2012; Menzerath, 1954). This principle has been applied to several systems – molecules, genes, music, and human language (Altmann, 1980; Baixeries, Hernández-Fernández, & Ferrer-i-Cancho, 2012; Boroda & Altmann, 1991; Shahzad, Mittenthal, & Caetano-Anollés, 2015). By applying this law to multiple levels of gelada vocal communication (i.e., sequences and wobbles) we show that vocal complexity does not necessary evolve haphazardly, but rather by following principles that balance complexity with brevity (Chapter 4). However, it is also important to note that Menzerath's law might apply to non-communication systems and may not be only a 'linguistic law'. Future work will be needed to determine the extent to which Menzerath's law applies to a broad range of information-based systems.

Future directions

In addition to providing new insights into research on the ultimate origins vocal complexity, the work described in this dissertation inspires several avenues for future research regarding the proximate origins of vocal complexity.

Ontogeny of vocal complexity

A growing body of evidence suggests that non-human animals have a rudimentary capacity to learn how to produce their vocalizations or the appropriate contexts in which to make them (Petkov & Jarvis, 2012). However, our understanding of the development of elaboration in combinations of sound is mostly limited to three lineages of birds (e.g., hummingbirds, parrots, and songbirds). These taxa exhibit high levels of vocal learning during critical periods, and their ability to effectively learn can be inhibited by disruptions to their development (Brumm, Zollinger, & Slater, 2009). These processes align closely with patterns of vocal learning in humans, who learn language fundamentals in the first year of life (Kuhl, 2004). The human ability to effectively produce spoken language is also sensitive to early life experiences (Harlaar, DeThorne, Smith, Betancourt, & Petrill, 2016; Hoff, 2006). In addition, a host of genes are thought to set up the human (and avian) brain to be able to learn complex communication (Fisher & Vernes, 2015; Petkov & Jarvis, 2012). FOXP2, for example, is a candidate gene that encodes a transcription factor that is involved in the construction of neural systems controlling orofacial coordination (Enard, 2011; Fisher & Scharff, 2009).

Developmental studies of the vocal complexity are rare outside of humans and birds. Our work suggests that geladas may be a useful species in which to make such observations. Given the paucity of studies demonstrating that primates learn their vocalizations, the null hypothesis is that the production of complex sequences are innate. Still, we know little about infant, juvenile, sub-adult, and bachelor male vocal behavior in geladas. If male geladas were to learn how to produce complex vocal sequences, then we would expect that males born to new leader males would develop more complex vocal abilities than males born to older leader males (based on results from Chapter 8). Even if the ability to produce complex vocal sequences is innate, gelada males may learn the appropriate contexts in which to produce these sequences while living in their natal units or in all-male bachelor groups. Another line of inquiry could be to look at the structure of candidate genes related to vocal capacity. Although FOXP2 is generally conserved across taxa, recent evidence suggests that mutated forms of FOXP2 exist in echolocating bats compared to non-echolocating bats (G. Li, Wang, Rossiter, Jones, & Zhang, 2007; Webb & Zhang, 2005). As in bats, comparing the structure of the candidate genes involved with vocal control in geladas to other mammals could illuminate potential patterns of divergence involving genetic mutations.

Mechanisms of vocal complexity

A wealth of literature describes the neural mechanisms underlying the production of vocal communication in vertebrates (Jürgens, 2002, 2009; Petkov & Jarvis, 2012; Simonyan, Horwitz, & Jarvis, 2012). These systems include cortico-stratial-thalamic loops that differ based on whether the vocalizations of interest are learned or innate. Most of the work investigating links between neural mechanisms and vocal complexity focus on songbird models (Ca & Devoogd, 2000; DeVoogd, 2004; J. Li, Zeng, Zhang, & Zuo, 2006; Markman et al., 2008; Schmidt, Moore, MacDougall-Shackleton, & MacDougall-Shackleton, 2013; Spencer, Buchanan, Leitner, Goldsmith, & Catchpole, 2005). A limitation of this avian work, however, is that there are major structural differences in mammalian and avian brains, and it can often be difficult to make inferences to mammalian neural systems.

Geladas would make for an exciting non-human primate model of the neural underpinnings of vocal complexity, except that invasive studies needed to answer these questions are not currently feasible. Instead, future studies trying to link gelada vocal complexity to proximate mechanisms will need to focus on hormones (e.g., androgens and glucocorticoids) that can be measured in fecal samples. In addition, a growing body of evidence suggests that rodents might be useful for exploring neural underpinnings of vocal complexity. Mice produce multi-component songs during affiliative interactions, and these songs are sometimes disrupted when altering components of the cortico-striatal-thalamic pathways (Burgdorf, Wood, Kroes, Moskal, & Panksepp, 2007; Fischer & Hammerschmidt, 2011; Lahvis, Alleva, & Scattoni, 2011; Wöhr, van Gaalen, &

Schwarting, 2015; Wright, Dobosiewicz, & Clarke, 2013). These neurobiological studies of rodent ultrasonic vocalizations (USVs) focus on rats and mice.

Another excellent study system is that of voles. Not only is there extreme variation in mating systems both within and between vole species, but one of the species – prairie voles (*Microtus orchogaster*) – form long-term male-female pair bonds. Like geladas, prairie voles exhibit male-biased production of diverse vocalizations during affiliative interactions (Getz, 1993; Lepri, Theodorides, & Wysocki, 1988; Ma, Resendez, & Aragona, 2014; Williams, Catania, & Carter, 1992). As a supplement to my dissertation research, I ran a series of experiments exploring the links between the dopamine system and vocal complexity (i.e., repertoire size) in prairie voles in Dr. Brandon Aragona's lab at the University of Michigan. These experiments reveal that stimulating specific types of dopamine receptors in the brain (i.e., D2-like receptors) elicit higher USV rate and repertoire size in males in response to exposure to females. D2-like receptors also promote the formation of pair bonds in prairie voles (Aragona et al., 2006; Aragona & Wang, 2009; Wang et al., 1999). Thus, these data suggest that the same mechanisms promoting social bonding also promote vocal complexity, at least in prairie voles. Additionally, the USV repertoire size of males when they first meet a female is positively associated with the level of dopamine (DA) transmission in the nucleus accumbens (NAc) shell (i.e., a primary component of the brain's reward and motivation systems) after establishing a pair bond with the same female. Previous research suggests that males in stronger pair bonds show greater DA transmission in the NAc shell (Aragona et al., 2006; Aragona & Wang, 2009). Thus, my findings suggest that a male's vocal complexity predicts his ability to form a 'pair-bonded brain'. I propose that the vole lineage will make for an excellent model system in which to simultaneously study the proximate and ultimate functions of vocal complexity.

References

Altmann, G. (1980). Prolegomena to Menzerath's law. *Glottometrika*, *2*, 1–10.

Aragona, B. J., Liu, Y., Yu, Y. J., Curtis, J. T., Detwiler, J. M., Insel, T. R., & Wang, Z. (2006). Nucleus accumbens dopamine differentially mediates the formation and maintenance of monogamous pair bonds. *Nature Neuroscience*, *9*, 133–9.

Aragona, B. J., & Wang, Z. (2009). Dopamine regulation of social choice in a monogamous rodent species. *Frontiers in Behavioral Neuroscience*, *3*, 15.

Baixeries, J., Hernández-Fernández, A., & Ferrer-i-Cancho, R. (2012). Random models of Menzerath-Altmann law in genomes. *BioSystems*, *107*, 167–73.

Berwick, R. C., Okanoya, K., Beckers, G. J. L., & Bolhuis, J. J. (2011). Songs to syntax: The linguistics of birdsong. *Trends in Cognitive Sciences*, *15*, 113–121.

Boë, L.-J., Berthommier, F., Legou, T., Captier, G., Kemp, C., Sawallis, T. R., … Fagot, J. (2017). Evidence of a vocalic proto-system in the baboon (*Papio papio*) suggests pre-hominin speech precursors. *PloS One*, *12*, e0169321.

Boroda, M. G., & Altmann, G. (1991). Menzerath's law in musical texts. In M. G. Boroda (Ed.), *Musikometrika 3* (pp. 1–13). Bochum: Brockmeyer.

Brumm, H., Zollinger, S. A., & Slater, P. J. B. (2009). Developmental stress affects song learning but not song complexity and vocal amplitude in zebra finches. *Behavioral Ecology and Sociobiology*, *63*, 1387–1395.

Burgdorf, J., Wood, P. L., Kroes, R. A., Moskal, J. R., & Panksepp, J. (2007). Neurobiology of 50-kHz ultrasonic vocalizations in rats: electrode mapping, lesion, and pharmacology studies. *Behavioural Brain Research*, *182*, 274–283.

Ca, D. C. A., & Devoogd, T. J. (2000). Greater song complexity is associated with augmented song system anatomy in zebra ® nches, *11*, 2339–2344.

Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., & Townsend, S. W. (2014). Language evolution: syntax before phonology? *Proceedings of the Royal Society Series B: Biological Sciences*, *281*, 20140263.

D'Arnoldi, C. T., Foulley, J.-L., & Ollivier, L. (1998). An overview of the Weitzman approach to diversity. *Genetics, Selection, Evolution*, *30*, 149–161.

DeVoogd, T. J. (2004). Neural constraints on the complexity of avian song. *Brain, Behavior and Evolution*, *63*, 221–32.

Enard, W. (2011). FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. *Current Opinion in Neurobiology*, *21*, 415–424.

Fischer, J., & Hammerschmidt, K. (2011). Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: Insights into the evolution of vocal communication. *Genes, Brain and Behavior*, *10*, 17–27.

Fisher, S. E., & Scharff, C. (2009). FOXP2 as a molecular window into speech and language. *Trends in Genetics*, *25*, 166–177.

Fisher, S. E., & Vernes, S. C. (2015). Genetics and the language sciences. *Annual Review of Linguistics*, *1*, 289–310.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, *4*, 258–267.

Fitch, W. T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.

Fitch, W. T., de Boer, B., Mathur, N., & Ghazanfar, A. A. (2016). Monkey vocal tracts are speech-ready. *Science Advances*, *2*, e1600723.

Getz, L. (1993). Social organization of the Prairie vole (Microtus ochrogaster). *Journal of Mammalogy*, *74*, 44–58.

Gustison, M. L., & Bergman, T. J. (In Review). Divergent acoustic properties of gelada and baboon vocalizations and the implications for the evoluation of human speech. *Journal of Language Evolution*.

Gustison, M. L., & Bergman, T. J. (2016). Vocal complexity influences female responses to gelada male calls. *Scientific Reports*, *6*, 19680.

Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (Theropithecus gelada) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *367*, 1847–59.

Gustison, M. L., Semple, S., Ferrer-i-Cancho, R., & Bergman, T. J. (2016). Gelada vocal sequences follow Menzerath's linguistic law. *Proceedings of the National Academy of Sciences*, *113*, E2750–E2758.

Harlaar, N., DeThorne, L. S., Smith, J. M., Betancourt, M. A., & Petrill, S. A. (2016). Longitudinal effects on early adolescent language: A twin study. *American Journal* *of Speech-Language Pathology*, *25*, 1–15.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.

Hauser, M. D., Yang, C., Berwick, R. C., Tattersall, I., Ryan, M. J., Watumull, J., … Lewontin, R. C. (2014). The mystery of language evolution. *Frontiers in Psychology*, *5*, 1–12.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*, 55–88.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews*, *26*, 235–258.

Jürgens, U. (2009). The neural control of vocalization in mammals: A review. *Journal of Voice*, *23*, 1–10.

Kempe, V., Puts, D. A., & Cárdenas, R. A. (2013). Masculine men articulate less clearly. *Human Nature*, *24*, 461–475.

Kershenbaum, A. (2014). Entropy rate as a measure of animal vocal complexity. *Bioacoustics*, *23*, 195–208.

Köhler, R. (2012). Quantitative syntax analysis. Berlin: De Gruyter Mouton.

Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews. Neuroscience*, *5*, 831–43.

Lahvis, G. P., Alleva, E., & Scattoni, M. L. (2011). Translating mouse vocalizations: Prosody and frequency modulation. *Genes, Brain and Behavior*, *10*, 4–16.

Lepri, J. J., Theodorides, M., & Wysocki, C. J. (1988). Ultrasonic vocalizations by adult prairie voles, Microtus ochrogaster. *Experientia*, *44*, 271–273.

Li, G., Wang, J., Rossiter, S. J., Jones, G., & Zhang, S. (2007). Accelerated FoxP2 evolution in echolocating bats. *PLoS ONE*, *2*. http://doi.org/10.1371/journal.pone.0000900

Li, J., Zeng, S. J., Zhang, X. W., & Zuo, M. X. (2006). The distribution of substance P and met-enkephalin in vocal control nuclei among oscine species and its relation to song complexity. *Behavioural Brain Research*, *172*, 202–11.

Lieberman, P. H., Klatt, D. H., & Wilson, W. H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science*, *164*, 1185–1187.

Luo, B., Jiang, T., Liu, Y., Wang, J., Lin, A., Wei, X., & Feng, J. (2013). Brevity is prevalent in bat short-range communication. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, *199*, 325–33.

Ma, S. T., Resendez, S. L., & Aragona, B. J. (2014). Sex differences in the influence of social context, salient social stimulation and amphetamine on ultrasonic vocalizations in prairie voles. *Integrative Zoology*, *9*, 280–93.

Manser, M. B. (2013). Semantic communication in vervet monkeys and other animals. *Animal Behaviour*, *86*, 491–496.

Markman, S., Leitner, S., Catchpole, C., Barnsley, S., Müller, C. T., Pascoe, D., & Buchanan, K. L. (2008). Pollutants increase song complexity and the volume of the brain area HVC in a songbird. *PloS One*, *3*, e1674.

Menzerath, P. (1954). *Die Architektonik des deutschen Wortschatzes*. Dümmler: Bonn.

Page, S. E. (2010). *Diversity and Complexity*. Princeton, New Jersey: Princeton University Press.

Petkov, C. I., & Jarvis, E. D. (2012). Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Frontiers in Evolutionary Neuroscience*, *4*, 1–24.

Poincaré, H. (1913). *The foundations of science: science and hypothesis, the value of science, science and method*. New York: The Science Press.

Popper, K. (1934). The logic of scientific discovery. Vienna: Verlag von Julius Springer.

Rosenberg, J., & Tunney, R. J. (2008). Human vocabulary use as display. *Evolutionary Psychology*, *6*, 538–549.

Schmidt, K. L., Moore, S. D., MacDougall-Shackleton, E. a., & MacDougall-Shackleton, S. a. (2013). Early-life stress affects song complexity, song learning and volume of the brain nucleus RA in adult male song sparrows. *Animal Behaviour*, *86*, 25–35.

Semple, S., Hsu, M. J., & Agoramoorthy, G. (2010). Efficiency of coding in macaque vocal communication. *Biology Letters*, *6*, 469–71.

Seyfarth, R. M., Cheney, D. L., Bergman, T. J., Fischer, J., Zuberbühler, K., & Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Animal Behaviour*, *80*, 3–8.

Shahzad, K., Mittenthal, J. E., & Caetano-Anollés, G. (2015). The organization of domains in proteins obeys Menzerath-Altmann's law of language. *BMC Systems Biology*, *9*, 44.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System's Technical Journal*, *27*, 379–423.

Simonyan, K., Horwitz, B., & Jarvis, E. D. (2012). Dopamine regulation of human speech and bird song: A critical review. *Brain and Language*, *122*, 142–150.

Spencer, K. a, Buchanan, K. L., Leitner, S., Goldsmith, A. R., & Catchpole, C. K. (2005). Parasites affect song complexity and neural development in a songbird. *Proceedings of the Royal Society B: Biological Sciences*, *272*, 2037–43.

Suzuki, K., Ikebuchi, M., Bischof, H.-J., & Okanoya, K. (2014). Behavioral and neural trade-offs between song complexity and stress reaction in a wild and a domesticated finch strain. *Neuroscience and Biobehavioral Reviews*, http://dx.doi.org/10.1016/j.neubiorev.2014.07.011.

Wang, Z., Yu, G., Cascio, C., Liu, Y., Gingrich, B., & Insel, T. R. (1999). Dopamine D2 receptor-mediated regulation of partner preferences in female prairie voles (Microtus ochrogaster): A mechanism for pair bonding? *Behavioral Neuroscience*, *113*, 602–611.

Webb, D. M., & Zhang, J. (2005). FoxP2 in song-learning birds and vocal-learning mammals. *Journal of Heredity*, *96*, 212–216.

Weitzman, M. L. (1992). On diversity A. *The Quarterly Journal of Economics*, *107*, 363–405.

Williams, J. R., Catania, K. C., & Carter, C. S. (1992). Development of partner preferences in female prairie voles (*Microtus ochrogaster*): the role of social and sexual experience. *Hormones and Behavior*, *26*, 339–49.

Wöhr, M., van Gaalen, M. M., & Schwarting, R. K. (2015). Affective communication in rodents. *Behavioural Pharmacology*, *26*, 506–521.

Wright, J. M., Dobosiewicz, M. R. S., & Clarke, P. B. S. (2013). The role of dopaminergic transmission through D1-like and D2-like receptors in amphetamine-induced rat ultrasonic vocalizations. *Psychopharmacology*, *225*, 853–68.