





Vocal Communication within the Genus *Chlorocebus*  
Insights into Mechanisms of Call Production and Call Perception

Dissertation

for the award of the degree

"Doctor of Philosophy"

Division of Mathematics and Natural Sciences  
of the Georg-August-Universität Göttingen

within the doctoral program Biology  
of the Georg-August University School of Science (GAUSS)

submitted by

**Tabitha Price**

from Hitchin, UK

Göttingen, 2013

### Thesis Committee

Prof. Dr. Julia Fischer, Cognitive Ethology Lab, German Primate Center GmbH

Prof. Dr. Peter Kappeler, Behavioral Ecology and Sociobiology Unit, German Primate Center GmbH

### Members of the Examination Board

Reviewer:

Prof. Dr. Julia Fischer

Second Reviewer:

Prof. Dr. Peter Kappeler

Additional Reviewer:

Prof. Dr. Stuart Semple, Department of Life Sciences, University of Roehampton

### Further members of the Examination Board

Prof. Dr. Eckhard Heymann, Behavioral Ecology and Sociobiology Unit, German Primate Center GmbH

Prof. Dr. Julia Ostner, Primates Social Evolution Group, Courant Research Centre Evolution of Social Behaviour

Prof. Dr. Dirk Semmann, Evolution of Cooperation and Prosocial Behaviour Group, Courant Research Centre Evolution of Social Behaviour

Prof. Dr. Ralf Heinrich, Cellular Neurobiology Department, Schwann-Schleiden Research Centre

**Date of the oral examination: 04.09.2013**

*“Language is the Rubicon which divides man From beast, and no animal will ever cross it . . . the science of language will yet enable us to withstand the extreme theories of the Darwinians, and to draw a hard and fast line between man and brute”*

*Max Müller 1862*

*“It is not the mere power of articulation that distinguishes man from other animals, for as everyone knows, parrots can talk; but it is his large power of connecting definite sounds with definite ideas, and this obviously depends on the development of the mental faculties”*

*Charles Darwin, 1871, The Descent of man*



# Table of Contents

|  |            |
|--|------------|
| <b>Summary</b> .....   | <b>i</b>   |
| <b>Zusammenfassung</b> .....   | <b>v</b>   |
| <b>Résumé</b> .....  | <b>ix</b>  |
| <b>Chapter 1. General Introduction</b> .....   | <b>1</b>   |
| 1.1 The comparative approach to communication .....  | 2          |
| 1.1.1 Co-operation or manipulation?.....   | 2          |
| 1.1.2 Form and function.....   | 3          |
| 1.1.3 A communicative origin of language?.....   | 6          |
| 1.2 The multi-component approach to language evolution .....   | 6          |
| 1.2.1 Vocal flexibility .....  | 8          |
| 1.2.2 Semantics .....  | 11         |
| 1.2.3 Pragmatics.....  | 15         |
| 1.3 African green monkeys: A model taxon .....   | 17         |
| 1.4 Study aims and approaches.....   | 18         |
| <b>Chapter 2. Limited geographic variation in the acoustic structure of and responses to adult male alarm barks of African green monkeys</b> ..... | <b>21</b>  |
| <b>Chapter 3. Vervets revisited: A quantitative analysis of alarm call structure and context specificity</b> .....                                 | <b>43</b>  |
| <b>Chapter 4. Meaning attribution in the West African green monkey: Influence of call type and context</b> .....                                   | <b>71</b>  |
| <b>Chapter 5. General Discussion</b> .....   | <b>91</b>  |
| 5.1 Proximate Mechanisms .....   | 91         |
| 5.1.1 Call production .....  | 91         |
| 5.1.2 Call perception .....  | 95         |
| 5.1.3 Appraisal theories of emotion.....   | 98         |
| 5.2 Ultimate Mechanisms.....   | 100        |
| 5.2.1 Evolution of alarm calls .....   | 100        |
| 5.2.2 Evolution of language.....   | 103        |
| 5.3 Conclusion and Outlook .....   | 105        |
| <b>References</b> .....  | <b>107</b> |
| <b>Acknowledgements</b> .....  | <b>131</b> |



## Summary

It is difficult to overestimate the influence of language in day to day life, and language is often proposed as the clearest feature to distinguish us, *Homo sapiens*, from all other animals. The origin of language has long been the subject of speculation and myth, and more recently, together with the development of Darwinian evolutionary concepts, the evolution of language has become the focus of multidisciplinary scientific investigation. The emerging picture is that the language faculty is a complex multi-component system, with the possibility of each subcomponent having arisen from a different evolutionary pathway. Comparative studies of communication and cognition in non-human animals have the potential to identify whether these subcomponents, or biological precursors to them, are shared with other species, offering insights into the communicative abilities of our pre-linguistic ancestors and the evolutionary precursors to language. To date there is no comprehensive list of what constitutes the language faculty; nevertheless vocal flexibility, semantic meaning and pragmatic inference are all fundamental characteristics of language and comparative studies can thus be used to shed light on how these particular features arose.

In this thesis I investigate vocal communication in the African green monkey (*Chlorocebus*), with especial focus on alarm calls. The African green monkey is a widespread genus consisting of six species, one of which is the vervet monkey (*C. pygerythrus*). Numerous studies have been carried out on vocal communication in the vervet, and their alarm calling system is the classic example of functional reference (that is, calls that function in a referential manner) in a non-human species. Nevertheless there are still many open questions regarding the mechanisms underlying call production and call perception in this species, and very little is known about vocal variability across the genus. In three inter-related studies I therefore explore questions of vocal flexibility with respect to the degree of control over vocal production, questions of semanticity concerning the relation between vocal signals and external stimuli, and questions of pragmatics pertaining to whether signal receivers incorporate contextual cues when responding to conspecific calls. The overall aim of these studies is to increase understanding of the proximate mechanisms and selective pressures that shape the production and perception of functionally referential alarm calls in the African green monkey, with the broader goal of offering insights into the emergence of vocal flexibility and meaning in language.

To investigate vocal flexibility in the adult male alarm bark, I carried out a study to assess geographic variability in vocal structure within and between African green monkey species, and the function of any structural differences. To do this I recorded alarm barks from male green monkeys (*C. sabaues*) in Senegal, and from male vervets (*C. p. pygerythrus*) in South Africa, and compared these to alarm barks made available to me from East African male vervets (*C. p. hilgerti*). I then carried out playback experiments in which the barks of an own-group male, an unknown conspecific male and unknown heterospecific (but congeneric) male were broadcast to adult male South African vervets to test whether they would distinguish between them. All barks demonstrated a similar overall call structure; nevertheless, inter-specific differences in fine-scale

spectral and temporal characteristics allowed green monkey barks to be distinguished from vervet barks with a high degree of accuracy. East African and South African vervet barks also exhibited spectral differences in call structure, although intra-specific differences were smaller than inter-specific differences. In response to playback experiments, males typically reacted most strongly to unknown conspecific barks, and showed behaviours typical of leopard-avoidance and male-male aggression. The findings of this study thus imply that the acoustic structure of male barks is innate, and that variation is likely related to phylogenetic differences. Furthermore, barks may have a dual function as both an alarm and display call, offering a potential explanation for the sexual dimorphism observed in *Chlorocebus* alarm calls.

The seminal finding that vervet monkeys produce predator-specific alarm calls that elicit predator-appropriate response behaviours led initially to the conclusion that these calls are semantic signals akin to human words. The inferences that can be drawn regarding a call's semantic value based only on the responses it evokes in others are, however, limited. In a second study, I revisit the vervet alarm calling system. Using recordings made available from the original studies together with additional calls I recorded myself, I carried out quantitative acoustic analyses to assess the degree to which calls are structurally discrete and context specific, both classic criteria for functional reference. These acoustic analyses reveal that at the level of the call element, vervet alarm calls are distinguishable with regards to the predator type eliciting calling, but that differences in spectral structure are rather graded; especially in the case of male calls. When compared with calls produced during social and aggressive interactions, female and male alarm call elements are less distinct. Importantly, temporal characteristics of call series distinguished between spectrally similar calls produced in different contexts. Thus at the level of the call element, vervet alarm calls constitute a graded system, especially in males, and context-specificity is reduced when calls from non-predator contexts are also considered. The adaptive value of distinguishable alarm calls in this genus, may have led to the production of temporally distinctive call series.

Green monkey females produce acoustically similar alarm chirps to snake and leopard models. In study three I 1) quantified acoustic differences between snake and leopard chirps, 2) used predator models to identify typical response behaviours to snakes and leopards, and 3) conducted playback experiments to assess whether subjects would respond to chirps with predator-typical response behaviours, and whether contextual cues in the form of pre-exposure to a leopard or snake model would influence these responses. Chirp calls exhibited graded differences between leopard and snake contexts; nevertheless, subjects responded to leopard chirps with leopard-appropriate behaviour irrespective of contextual cues. Interestingly, priming condition did not have an effect on the subjects' immediate response, but, together with call type, did influence subsequent behaviour.

In conclusion, the observed variation in vocal structure across species of the African green monkey is likely related to phylogenetic differences, with more flexibility in temporal than spectral characteristics. In keeping with findings from all other primate taxa to date, this is

indicative of a largely innate vocal production process with limited control over call structure. At the level of the call element, vervet alarm calls show considerable intergradation, both among different alarm call types and between alarm and non-alarm calls, a finding in line with neurological studies that suggest that structural features of primate calls are strongly influenced by the caller's motivational state. Considering call perception, acoustic and contextual cues influence green monkey responses to a graded alarm call. Taken together, the findings of these studies suggest that functionally referential signalling systems provide little insight, on the side of the signaller, into the evolution of flexible and arbitrary symbolic communication. More relevant questions are how cognitive and emotional factors interact to influence call production, and how contextual cues are integrated into processes of call perception in non-human animals. Such comparative research is highly relevant to studies of emotional prosody and pragmatic inference in humans, with implications for a better understanding of how language evolved from the innate and emotionally grounded communication system of our pre-linguistic ancestors.



## Zusammenfassung

Der Einfluss der Sprache auf das tägliche Leben kann kaum hoch genug eingeschätzt werden, und die Sprache wird oft auch als das eindeutigste Merkmal angesehen, das uns, *Homo sapiens*, von allen anderen Tierarten unterscheidet. Der Ursprung der Sprache war lange Objekt von Spekulationen und Mythen. Mit dem Aufkommen von Darwins Evolutionstheorie rückte die Evolution der Sprache in den Fokus disziplinübergreifender wissenschaftlicher Untersuchungen. Das sich herauskristallisierende Bild zeigt die Sprache als ein komplexes, aus mehreren Komponenten bestehendes System, wobei möglicherweise jede Einzelkomponente einem eigenen evolutionären Pfad gefolgt ist. Vergleichende Untersuchungen der Kommunikation und Kognition von Tieren ermöglichen es festzustellen, ob diese Einzelkomponenten oder ihre biologischen Vorstufen auch bei Tieren anzutreffen sind. Hierdurch ist es möglich, Erkenntnisse über die kommunikativen Fähigkeiten unserer prä-linguistischen Vorfahren und die evolutionären Vorstufen unserer Sprache zu gewinnen. Bis heute existiert keine Klarheit darüber, was Sprache eigentlich auszeichnet. Gleichwohl ist es unstrittig, dass vokale Flexibilität, semantische Bedeutung und pragmatische Rückschlüsse grundlegende Aspekte von Sprache sind. Folglich erlauben uns vergleichende Studien ein besseres Verständnis über die Entstehung dieser Fähigkeiten zu gewinnen.

In meiner Arbeit untersuchte ich die vokale Kommunikation von Grünen Meerkatzen (*Chlorocebus*), unter besonderer Berücksichtigung von Alarmrufen. Grüne Meerkatzen sind eine weitverbreitete Gattung, die sich aus sechs Arten zusammensetzt. Eine dieser Arten ist die südliche Grünmeerkatze (*C. pygerythrus hilgerti*) aus Ostafrika. In mehreren Studien wurde die vokale Kommunikation dieser Grünmeerkatzenart analysiert, und ihr Alarmrufsystem gilt heute als klassisches Beispiel für funktional referenzielle Kommunikation bei nicht-menschlichen Arten. Trotz dieser Untersuchungen blieben viele Fragen über die Mechanismen der Lautproduktion und Perzeption unbeantwortet. Darüber hinaus liegen nur sehr wenige Informationen zur vokalen Flexibilität innerhalb der Gattung vor. In drei thematisch verbundenen Studien habe ich mich deshalb mit den folgenden Fragen beschäftigt: 1) Mit Blick auf die vokale Flexibilität untersuchte ich, inwieweit die Tiere Kontrolle über ihre Lautäußerungen haben; 2) bezüglich der Semantik betrachtete ich den Zusammenhang von externen Reizen und akustischen Signalen, und 3) beschäftigte ich mich hinsichtlich der Pragmatik mit der Frage, ob die Empfänger von Signalen bei der Reaktion auf die Rufe von Artgenossen den jeweiligen situativen Kontext berücksichtigen. Das übergreifende Ziel meiner Studien ist es unser Verständnis der proximalen Mechanismen und Selektionsdrücke, die die Produktion und Wahrnehmung von funktional referentiellen Alarmrufen in Grünen Meerkatzen beeinflussen, zu erweitern. Dies kann gleichzeitig tiefere Einblicke in die Evolution vokaler Flexibilität und die Bedeutung in der Sprache liefern.

Um die vokale Flexibilität eines bestimmten Typs von Alarmrufen von Männchen („Barks“) zu untersuchen, führte ich eine Studie zur geographischen Variabilität in der akustischen Struktur dieser Rufe innerhalb und zwischen verschiedenen Arten von Grünen Meerkatzen

durch, wobei ich auch mögliche Funktionen der strukturellen Unterschiede zwischen den Rufen berücksichtigte. Hierfür nahm ich „Barks“ von männlichen westafrikanischen Grünmeerkatzen (*C. sabaeus*) im Senegal und von männlichen südlichen Grünmeerkatzen (*C. p. pygerythrus*) in Südafrika auf, und verglich diese mit Rufen von männlichen südlichen Grünmeerkatzen (*C. p. hilgerti*) aus Ostafrika, die mir zur Verfügung gestellt wurden. Im Anschluss daran führte ich Playback-Experimente durch, bei denen ich südafrikanischen Grünmeerkatzen die Rufe eines bekannten Männchens aus ihrer Gruppe, eines unbekanntes Artgenossen, und eines Männchens einer anderen Art von Grünmeerkatzen vorspielte, um zu untersuchen, ob sie zwischen diesen verschiedenen Rufen unterscheiden würden. Alle „Barks“ wiesen eine ähnliche Lautstruktur auf, jedoch gab es auf einer feineren Analyseskala zwischenartliche Unterschiede in den spektralen und zeitlichen Charakteristika der Rufe. Diese erlaubten es, die „Barks“ von westafrikanischen und anderen Grünmeerkatzen mit hoher Sicherheit zu unterscheiden. Auch die Struktur der „Barks“ von Grünmeerkatzen aus Ost- und Südafrika zeigten spektrale Unterschiede, wenn auch geringer als zwischen westlichen und anderen Grünmeerkatzenarten. In den Playback-Experimenten zeigte sich, dass die Männchen normalerweise am stärksten auf die Rufe von unbekanntes Artgenossen reagierten. Hierbei zeigten sie insbesondere Verhaltensweisen, die typischerweise als Reaktion auf Leoparden oder bei aggressiven Interaktionen zwischen Männchen auftreten. Diese Ergebnisse deuten darauf hin, dass die akustische Struktur der Rufe angeboren ist und dass bestehende Unterschiede vermutlich mit phylogenetischen Unterschieden zusammenhängen. Weiterhin deuten sie darauf hin, dass „Barks“ sowohl als Alarm- als auch „Display“- Rufe fungieren. Dies könnte auch den beobachteten Geschlechtsunterschied in den Alarmrufen von *Chlorocebus* erklären.

Die bahnbrechende Entdeckung, dass ostafrikanische Grünmeerkatzen raubfeindspezifische Alarmrufe haben, die adäquates, raubfeindspezifisches Verhalten auslösen, führte ursprünglich zu dem Schluss, dass diese Rufe semantische Signale seien, ähnlich Worten in der menschlichen Sprache. Die Rückschlüsse, die man an Hand der Reaktion auf einen Ruf über die semantische Bedeutung dieses Rufes ziehen kann, sind jedoch begrenzt. In meiner zweiten Studie analysierte ich daher detailliert das Alarmrufsystem der ostafrikanischen Grünmeerkatzen. Hierbei konnte ich sowohl auf die mir zur Verfügung gestellten Lautaufnahmen der ursprünglichen Studie, als auch auf zusätzliche, von mir aufgenommene Rufe zurückgreifen. Ich führte eine quantitative akustische Analyse durch um abzuschätzen, inwieweit die Rufe einerseits strukturell diskret sind und andererseits kontextspezifisch eingesetzt werden. Sowohl strukturelle Unterschiede als auch Kontextspezifität sind klassische Kriterien für funktionale Referenz. Die Analysen zeigen, dass die Alarmrufelemente raubfeindspezifisch sind, dass aber die Unterschiede der Spektralstruktur insbesondere bei Rufen von Männchen eher graduell sind. Anders als bei Rufen, die in sozialen und aggressiven Interaktionen geäußert werden, sind die Alarmrufe von Männchen und Weibchen weniger distinkt. Wichtig ist, dass Rufe, die in verschiedenen Kontexten auftraten und in ihrer Spektralstruktur ähnlich waren, sich in ihren zeitlichen Merkmalen unterschieden. Daraus ergibt sich, dass auf der Ebene der Rufelemente die Alarmrufe, insbesondere von männlichen ostafrikanischen Grünmeerkatzen, ein „Graded“-

System darstellen. Die Kontextspezifität ist reduziert, wenn man auch Situationen berücksichtigt in denen keine Raubfeinde anwesend sind. Der adaptive Wert unterscheidbarer Alarmrufe könnte zur Produktion von Alarmrufreihen mit unterschiedlichen zeitlichen Mustern geführt haben.

Weibliche westafrikanische Grünmeerkatzen produzieren akustisch ähnliche Alarm-„Chirps“, wenn sie Schlangen- oder Leopardentrappen sehen. In meiner dritten Studie habe ich daher 1) akustische Unterschiede der „Chirps“ untersucht, die als Reaktion auf Schlangen und Leoparden geäußert wurden, 2) Raubfeindattrappen eingesetzt, um typische Reaktionen auf Schlangen und Leoparden zu identifizieren, und 3) Playback-Experimente eingesetzt um festzustellen, ob die Tiere auf „Chirps“ mit raubfeindspezifischen Verhaltensweisen reagieren, und ob sie hierbei kontextuelle Information berücksichtigen, die auf einer vorangegangenen Konfrontation mit einer Schlangen- oder einer Leopardentrappe basieren. Bei diesem spezifischen Alarmruf fand ich graduelle Unterschiede zwischen Rufen, die als Reaktion auf eine Schlange oder einen Leopard geäußert wurden. Trotzdem zeigten die Tiere entsprechend leopardenspezifische Reaktionen, wenn sie einen Leopard-Warnruf gehört hatten, unabhängig von dem jeweiligen Kontext. Interessanterweise hatte die vorangegangene Konfrontation mit einem Raubfeindmodell keinen Einfluss auf ihr direktes Verhalten, hatte aber, in Kombination mit dem jeweiligen Ruftyp, einen Einfluss auf ihr späteres Verhalten.

Abschließend lässt sich sagen, dass die beobachtete Variabilität in der vokalen Struktur der Rufe zwischen Arten von Grünen Meerkatzen vermutlich mit phylogenetischen Unterschieden zusammenhängt. Flexibilität ergibt sich eher in temporalen als in strukturellen Aspekten der Rufe. Dies deckt sich mit Ergebnissen, die bei anderen Primatenarten gefunden wurden, und deutet auf einen größtenteils angeborenen Prozess der Rufgenerierung und einer eingeschränkte Kontrolle über die Struktur der Rufe hin. Auf der Ebene von Rufelementen zeigen die Rufe von ostafrikanischen Grünmeerkatzen beträchtliche graduelle Überschneidungen, sowohl zwischen den verschiedenen Typen von Alarmrufen als auch zwischen Alarmrufen und Nicht-Alarmrufen. Neurologische Studien legen zudem nahe, dass die strukturellen Komponenten der Rufe von Primaten stark von der Motivation der Rufer beeinflusst werden. Bei der Wahrnehmung der Rufe dieses „Graded“-Alarmrufsystems haben sowohl akustische als auch kontextuelle Faktoren einen Einfluss auf die Reaktion von der Meerkatzen. Zusammenfassend legen die Ergebnisse den Schluss nahe, dass funktional referentielle Signalsysteme hinsichtlich der Produzenten der Signale wenig Aufschluss über die Evolution flexibler und arbiträrer symbolischer Kommunikation geben. Von größerer Bedeutung sind Fragen nach dem Zusammenspiel kognitiver und emotionaler Faktoren bei der Rufproduktion, und wie kontextuelle Information die Wahrnehmung von Rufen bei Tieren beeinflusst. Interspezifische vergleichende Forschung ist von großer Bedeutung für die Untersuchung von menschlicher emotionaler Prosodie und inferentieller Pragmatik und für ein besseres Verständnis der Evolution von Sprache aus angeborenen und emotions-basierenden Kommunikationssystemen unserer prä-linguistischen Vorfahren unabdingbar.



## Résumé

L'importance de l'influence du langage sur la vie de tous les jours ne peut être sous-estimée, et bien souvent le langage est perçu comme l'élément qui marque le plus clairement la différence entre nous, *Homo sapiens*, et les autres animaux. L'origine du langage est depuis longtemps un sujet de spéculation et de mythe, et plus récemment avec le développement des concepts évolutionnaires darwiniens, l'évolution du langage est devenue le centre d'intérêt de recherches scientifiques multidisciplinaires. L'image qui en émerge est que la faculté du langage est un système complexe aux composantes multiples et qu'il est possible que chacune de ses sous-composantes soit issue d'une trajectoire évolutive distincte. Les études comparatives sur la communication et la connaissance chez les animaux non-humains permettent d'identifier si ces sous-composantes ou leurs précurseurs biologiques sont communs à d'autres espèces, offrant ainsi une meilleure perception des capacités de communication de nos ancêtres prélinguistiques et des précurseurs évolutifs du langage. Jusqu'à présent il n'y a pas de liste exhaustive de ce qui constitue la faculté du langage ; néanmoins, la flexibilité vocale, la signification sémantique et l'inférence pragmatique constituent toutes des caractéristiques fondamentales du langage et les études comparatives peuvent ainsi servir à élucider comment ces traits en particuliers sont apparus.

Dans cette thèse, j'étudie la communication vocale du *Chlorocebus*, avec une emphase toute particulière sur les cris d'alarme. *Chlorocebus* est un genre largement répandu et comprend six espèces dont l'une est le vervet (*C. pygerythrus*). De nombreuses études ont été faites sur la communication chez les vervets et leur système de cris d'alarme est un exemple classique de référence fonctionnelle (c'est-à-dire, des cris qui fonctionnent d'une façon référentielle) chez une espèce non-humaine. Néanmoins, de nombreuses questions restent ouvertes sur les mécanismes sous-jacents à la production et à la perception des cris chez cette espèce et les connaissances sur les variations vocales au sein du genre *Chlorocebus* restent très limitées. Dans trois études interconnectées, j'explore donc des questions relatives à la flexibilité vocale en considérant le degré de contrôle sur la production vocale, des questions de sémantique sur la relation entre les signaux vocaux et les stimuli externes, et des questions d'ordre purement pragmatique à savoir si les receveurs de signaux incorporent des indicateurs contextuels dans leurs réponses aux cris des conspécifiques. L'objectif général de ces études est de mieux comprendre les mécanismes proximaux et les pressions sélectives qui conduisent à la production et à la perception de "functionally referential" cris d'alarme chez le *Chlorocebus*, avec comme objectif plus global de mieux comprendre l'émergence de la flexibilité vocale et de la signification dans le langage.

Pour étudier la flexibilité vocale dans les cris de type « bark » des mâles adultes, j'ai analysé la variabilité géographique dans la structure vocale entre les différentes espèces de *Chlorocebus* et le rôle des différences structurelles rencontrées. Pour cela, j'ai enregistré les cris d'alarme de singes verts mâles (*C. sabaes*) au Sénégal, et ceux de vervets mâles (*C. p. pygerythrus*) en Afrique du Sud et les ai comparés aux cris d'alarme qui m'ont été fournis de vervets mâles

d'Afrique orientale (*C. p. hilgerti*). Ensuite, j'ai fait des expériences dans lesquelles les cris d'un mâle du même groupe, d'un mâle de la même espèce et d'un mâle d'une autre espèce non identifiée (mais du même genre) ont été diffusés à des vervets d'Afrique du Sud pour établir si ils pouvaient les distinguer. Tous ces cris avaient une structure globale similaire; néanmoins, des différences interspécifiques de moindre échelle entre les caractéristiques spectrales et temporelles ont permis de distinguer avec une grande précision les cris des singes verts de ceux des vervets. Le cri des vervets d'Afrique Orientale et d'Afrique du Sud démontrent aussi des différences spectrales dans leur structure. Néanmoins les différences intraspécifiques étaient de moindre envergure que les différences interspécifique. En réponse aux expériences de rediffusion des cris, les mâles réagirent le plus fortement aux « bark » des conspécifiques inconnus et affichèrent des réactions comportementales caractéristiques de fuite face à un léopard et d'agression entre mâles. Les résultats de cette étude suggèrent donc que la structure acoustique des cris des mâles est innée et que les variations sont probablement dues à des différences phylogénétiques. De plus, les « barks » peuvent avoir une double fonction de cris d'alarme et de « display », offrant ainsi une explication plausible au dimorphisme sexuel observé dans les cris d'alarme du *Chlorocebus*.

La découverte fondatrice que les singes vervet produisent des cris d'alarme spécifique à chaque prédateur qui génèrent en réponse un comportement approprié a initialement conduit à la conclusion que ces appels sont des signaux sémantiques semblables à des mots humains. La valeur sémantique d'un cri basée seulement sur les réponses qu'il génère chez d'autres n'amène cependant qu'à des déductions limitées. Dans la deuxième étude, je réexamine le système de cris d'alarme des vervets. En utilisant des enregistrements provenant de la première étude, enrichie d'autres enregistrés par moi-même, j'ai effectué des analyses acoustiques quantitatives pour évaluer à quel degré ces appels sont structurellement distincts et spécifiques au contexte, deux critères classiques de la référence fonctionnelle. Ces analyses acoustiques révèlent qu'au niveau des éléments du cri, on peut distinguer les cris en fonction du prédateur qui en est la cause, mais que les différences dans la structure spectrale sont plus graduelles, surtout pour les cris des mâles. En comparaison avec les cris pendant les activités sociales et agressives, les éléments des cris d'alarme des mâles et des femelles sont moins distincts. Plus important encore, les caractéristiques temporelles des séries de cris étaient différenciables même entre des cris spectralement similaires produits dans des contextes différents. Ainsi, au niveau des éléments, les cris d'alarme des vervets constituent un système graduel, surtout chez les mâles et la spécificité de contexte se trouve réduite quand seul les cris produits dans un contexte dépourvu de prédation sont considérés. La valeur adaptative de cris d'alarme distinctifs dans le genre *Chlorocebus* peut avoir conduit à la production de séries de cris temporellement distincts.

Les singes vert femelles produisent des cris d'alarme « Chirps » acoustiquement similaires face à des modèles de serpent et de léopard. Dans la troisième étude, j'ai 1) quantifié les différences acoustiques entre les « Chirps » face au serpent et face au léopard, 2) identifié des réponses comportementales caractéristiques en réponse à un serpent et à un léopard en me

servant de modèles de ces prédateurs et 3) mené des expériences de rediffusion pour évaluer si les sujets réagiraient aux « Chirps » par un comportement caractéristique à une réponse face à un prédateur et si les signaux contextuels sous la forme de pré-exposition à un modèle de léopard ou de serpent pourraient influencer ces réactions. Les « Chirps » affichèrent des différences graduelles entre le contexte du singe et du léopard : cependant, les sujets répondirent aux « Chirps » face au léopard avec un comportement approprié sans tenir compte des signaux contextuels. Il est intéressant de noter que les pré-conditionnements n'ont pas affecté immédiatement le comportement des sujets, mais plus tard quand le type de cri a été ajouté.

En conclusion, la variation observée dans la structure vocale entre les différentes espèces de *Chlorocebus* est probablement liée à des différences phylogénétiques, avec une plus grande flexibilité des caractéristiques temporelles que des caractéristiques spectrales. À la lumière des connaissances actuelles sur les autres taxa de primates, cela indique un procédé de production vocale principalement inné avec un contrôle limité de la structure des cris. Au niveau des éléments du cri, les cris d'alarme des vervets montrent une considérable inter-gradation à la fois entre les différents types de cris d'alarme et entre les cris d'alarme et les cris produits en dehors de contextes liés à la menace de prédation, une conclusion en conformité avec les études neurologiques qui suggèrent que les traits structurels du cri des primates sont fortement influencés par l'état de motivation de l'animal. En ce qui concerne la perception du cri, les signaux acoustiques et contextuels influencent la réaction des *Chlorocebus* à un cri d'alarme graduel. Dans leur ensemble, les conclusions de ces études suggèrent que les systèmes signalétiques fonctionnellement référentiels ne fournissent que peu d'informations, du côté du signaleur, sur l'évolution d'une communication symbolique flexible et arbitraire. D'autres questions encore plus importantes restent à élucider : comment les facteurs cognitifs et émotionnels interagissent pour influencer la production des cris et comment les signaux contextuels sont-ils intégrés dans les procédés de perception des cris chez les animaux non-humains. Une telle recherche comparative serait très pertinente pour les études de prosodie émotionnelle et d'inférence pragmatique chez les humains, et pourrait permettre de mieux comprendre comment le langage a évolué à partir du système de communication innée et sans fondement émotionnel de nos ancêtres prélinguistiques.



# Chapter 1

## General Introduction

Language is a feature that distinguishes us, *Homo sapiens*, from all other extant species (Hauser et al. 2002). An evolutionary transition from primate-like calls to speech has been suggested as "the decisive step in the origin of specifically human society" (Maynard Smith and Szathmáry 1995, p12), and the evolving language faculty has been proposed as the basis from which all other uniquely human accomplishments developed (Snowdon 2004). This has led to the claim that:

*"The evolution of human language is thus one of the most significant and interesting evolutionary events that has occurred in the last 5-10 million years, and indeed during the history of life on Earth" (Fitch 2010, p.1)*

Nonetheless, very little is known about the origin and evolution of language, and it has been proposed that questions of language evolution may constitute the most difficult problem in science (Christiansen and Kirby 2003). Indeed some people, like Evolutionary biologist Richard Lewontin, doubt whether questions such as these can ever be answered.

*"Form and even behavior may leave fossil remains, but forces like natural selection do not. It might be interesting to know how cognition (whatever that is) arose and spread and changed, but we cannot know. Tough luck." (Lewontin, 1998, p.130)*

Certainly before considering such questions it is important to make clear that, as used here (and throughout this thesis), the terms language, and language evolution, refer to the system employed by humans to express thoughts, and how this system evolved, not to the study of cultural change within and between language types such as English or German. In addition, it is necessary to distinguish between language, as so defined, and speech, which is the verbal expression of language; and to clarify that the term language faculty is used to place emphasis on the biological capacity that underlies language and allows it to develop.

On the basis of these definitions, I begin this chapter by discussing how a comparative approach can offer insights into the selective forces and evolutionary constraints shaping communication systems, and consider the implications of these for an understanding of language origins. Following on from this, I introduce the "multi-component approach to language" (Fitch 2010), a framework within which comparative studies of communication and

cognition in non-human animals (hereafter animals) can offer insights into how different subcomponents of the language faculty evolved. Applying this framework I focus on three language components that are prominent in this thesis; vocal flexibility, semantics and pragmatics. For each of these three topics I discuss whether communicative or cognitive parallels or precursors are present in animals, and towards the end of the chapter I explain why the African green monkey (*Chlorocebus*) provides an excellent model taxon for furthering our understanding in these areas. Finally, I outline the overall aim of my thesis and the objectives and approaches of each study undertaken to achieve this aim.

### **1.1 The comparative approach to communication**

The comparative approach of Evolutionary Biology provides a powerful tool allowing researchers to identify species that share a specific trait, and, based on their evolutionary relationship, to determine the phylogenetic origin and potential function of this trait (Ridley 2004). But applying the comparative method to language is problematic because there is little to directly compare it with; there are no other living human species, no non-human communication system (that we know of) comes close to the complexity of language, and fossil evidence can provide few insights into the vocal behaviour of early hominins (Tincoff and Hauser 2009). Nevertheless, following the argument of Pinker and Bloom as laid out in their seminal paper, *Natural Language and Natural Selection* (Pinker and Bloom 1990), language is an adaptation that evolved through natural selection as a means of communication. Thus comparative studies of animal communication systems are integral to an understanding of the selective forces shaping communicative interactions and signal structure, and the extent to which these can explain the form and function of language today. More specifically, the primary function of modern human language is the co-operative exchange of cheap and reliable information from a speaker to a listener (Harley 2001). How does this differ from what we know of animal communication systems?

#### **1.1.1 Co-operation or manipulation?**

According to Dawkins and Krebs, the "classical ethological" perspective claimed of animal communication that "it is to the advantage of both parties that signals should be efficient, unambiguous and informative" (Dawkins and Krebs 1978, p. 289). Indeed, many more recent definitions of animal communication are based upon a similarly informationalist view. For instance, Bradbury and Vehrencamp state that communication is the "provision of information

from a sender to a receiver" (1998, p. 2). Dawkins and Krebs (1978) opposed this perspective on the logical basis that the evolution of cooperative behaviour is not an expected outcome of natural selection; rather natural selection should bring about individuals that produce signals in such a way as to manipulate the behaviour of signal receivers to their own advantage. As communication requires not only that signals are produced, but also that they elicit responses in signal receivers (Hinde 1981), Krebs and Dawkins later expanded their original argument to propose that selection should also act on receivers to become "mind-readers" in the sense that they would benefit from using signals to predict the signaller's subsequent actions (Krebs and Dawkins 1984). The outcome of this is that, at an ultimate level, animal communication is best seen not as a cooperative act, but as an evolutionary stable strategy (ESS, Maynard Smith 1978) within which both signaller and receiver must benefit on average; leading to the proposal that communication is the completion of a signalling act, within which both the production of and response to a signal evolved because of the effect the signal has on the receiver (Maynard Smith and Harper 1995; Scott-Phillips 2008). For an alternative view that does not require responses to have evolved for that purpose, see Fischer (2011). Importantly both of these "adaptationist" views of communication do not preclude that information plays a role in the proximate mechanisms underlying call perception (Fischer 2011), a point that will be re-visited when discussing semantics in section 1.2.2. That signallers and the mechanisms underlying call production should be considered separately from receivers and the mechanisms underlying call perception is an important concept in studies of animal communication (Seyfarth and Cheney 2003b; Wheeler et al. 2011), and one which will be a recurring theme throughout this thesis. In the next section I discuss the adaptive significance of signal form with respect to the selective forces imposed on manipulative signallers by mind-reading receivers and the environment in which they live.

### **1.1.2 Form and function**

The form of animal signals has been described as being influenced by two functional requirements, "strategic design", which is necessary for signal receivers to benefit from responding, and "tactical design", which is necessary for the signal to reach the receiver and elicit a response (Guilford and Dawkins 1991). Many of the design features of animal signals are shaped by these evolutionary constraints and I focus in more detail on each in turn within the following paragraphs.

### Strategic design

Strategic design is related to the concept that for receivers to benefit from responding to a signal, the signal must be "honest". An honest signal in this sense does not infer conscious or intentional "honesty" (Fitch 2010), but rather a signal whose structure co-varies reliably with an attribute of the signaller or environment (Fitch 2004). How such signals arise depends, in part, on how the evolutionary interests of signallers and receivers coincide (Searcy and Nowicki 2010).

In contexts within which the interests of signallers and receivers conflict or diverge, such as during aggressive interactions or mate choice, signals may be honest because they incur a cost to produce so that the degree of signal expression is proportional to caller fitness ("handicap trait", Zahavi 1975; 1977; Maynard Smith 1985; Grafen 1990), or because they cannot be faked because they are tied to the caller's motivational state ("motivational-structural" rules, Morton 1975) or are physically constrained ("indices of quality", Maynard Smith and Harper 1995). This physical constraint arises from the indexical relationship that exists between the spectral characteristics of many animal signals and signaller size, whereby the tension and length of vocal cords within the larynx determines their rate of oscillation and thus the fundamental frequency and harmonics of a sound (Fitch and Hauser 1995), whilst the length of the supra-laryngeal vocal tract is the primary determinant of a sound's resonant frequencies frequency (Fitch 2000). The roars of male red deer (*Cervus elaphus*), and the grunts and wahoos of baboons (*Papio*) constitute examples of both costly and indexical signals, with the anatomically constrained spectral properties of calls providing a reliable cue to body size (Reby and McComb 2003; Pfefferle and Fischer 2006) and the cost of a high calling rate providing a cue to caller condition (Albon and Clutton-Brock 1979; Kitchen et al. 2003; Fischer et al. 2004).

When the interests of signallers and receivers overlap, or when both parties rank interaction outcomes in the same order, signal reliability need not depend on physical constraints or costs of production (Maynard Smith and Harper 1995; Searcy and Nowicki 2010), and "cheap talk" can evolve. The Philip Sidney game (Maynard Smith 1991) provides a mathematical model in support of this, and an empirical example is when signalling stability is the result of kin selection, whereby the cost of signal production is small relative to the benefits inferred on related individuals (Hamilton's rule of inclusive fitness; Hamilton 1964). Some alarm calls (e.g. those of Belding's ground squirrels, *Urocitellus beldingi*; Sherman 1977) meet these assumptions by alerting relatives at small cost to the receiver. It has also been suggested that low-cost signals may evolve in situations when both parties benefit from coordinating their actions, or when individuals meet repeatedly and modify their responses on the basis of past interactions, see Silk and colleagues (2000) for further description and an empirical example.

### **Tactical design**

The tactical design or efficiency of a signal can be thought of as depending on signal design features, which are shaped in part by the functional requirements of a call (Owren and Rendall 2001; Semple and Higham 2013) in combination with constraints imposed by the signal receiver's sensitivity (Guilford and Dawkins 1991), as well as their identity and location (Hockett and Altmann 1968). One early evolutionary explanation for the distinct forms of close-contact signals within a species was put forward by Darwin in his “principle of antithesis” (Darwin 1872), within which he proposed that selection should act on close-contact calls to create a divergence of signals communicating conflicting signaller motivational states in order to prevent receivers from confusing them. This theory was expanded on by Morton (1977; 1982) to develop the motivational-structural rules mentioned above, with which he hypothesised that harsh low frequency calls (giving an impression of larger body size) would be related to aggressive motivational states, whilst tonal higher frequency calls (indicative of a smaller individual) would be related to submissive or fearful states. General support for a relationship between motivational state and call frequency (but not tonality) was provided by Hauser's (1993a) broad study across primate taxa. Macedonia and Evans (1993) also proposed that call structure should diverge in situations requiring incompatible responses, but they differed from Darwin and Morton in that they focused on alarm calls and suggested that call structure in this case would be related to the external stimulus (predator type) rather than the caller's internal state. This point is central to discussions of semantics and will be described in more detail in section 1.2.2.

Regarding structural variation within and between call types, Marler (1975; 1976) hypothesised that signallers should produce calls of a discrete structure (no intermediates between call types) when communicating over long distances or within closed habitats because the efficiency of the signal in these contexts would depend on the acoustic cues alone. In contrast signallers might be more likely to produce calls with a graded structure (continuous variation between call types) when communicating in close proximity or in open habitats when receivers can also take contextual cues from the signaller and environment into account. Guilford and Dawkins (1991) also stressed the importance of environmental constraints, and the “acoustic adaptation hypothesis” (Morton 1975) describes how differences in the acoustic environment (habitat type and/or ambient noise) can influence call structure so as to optimise transmission under the local conditions. For example, compared to species living in open habitats, forest-species tend to produce lower frequency calls with temporal features that are less sensitive to reverberation effects (Chappuis 1971; Morton 1975; Wiley 1991). It is important to note, however, that such constraints act only on calls that function in long-distance communication (Slabbekoorn and Smith 2002; Schneider et al. 2008).

Combining concepts of strategic and tactical design, predictions can be formulated regarding the function-dependent optimal structure of calls (Bradbury and Vehrencamp 1998). Empirical support that such hypotheses can be applied across a range of species comes from studies of flee and assembly alarm calls in birds and non-human primates (hereafter primates). Flee alarm calls that function to alert nearby conspecifics of urgent danger whilst remaining hard to locate tend to be short, tonal and high-frequency calls, whilst assembly alarm calls that function to assemble conspecifics from a wide area tend to be loud and exhibit a broadband structure (Marler 1955; 1968).

### **1.1.3 A communicative origin of language?**

At the beginning of the chapter, I described language as the co-operative exchange of cheap and reliable information from a speaker to a listener and asked how this differs from animal communication systems. The above paragraphs discuss the selective forces and evolutionary constraints shaping animal communication, and in doing so highlight some major differences between these communication systems and language. To constitute an evolutionary stable strategy, animal communication in many cases is either costly or else the structure of vocalisations is constrained by the callers physiology or motivational state. Call form is further related to maximising efficacy with regards to transmission within the environment and bringing about call function. Discussions of how language overcame or avoided these limitations have led to the suggestion that language need not have evolved directly from the communication system of our last common ancestor with chimpanzees (Fitch 2010). Rather, to understand how language circumvented these constraints it is necessary to tease apart the proximate mechanisms underlying the acquisition and use of language and to ask whether any of these mechanisms, or precursors of these mechanisms, are shared with other animals. The "multi-component approach to language" (Fitch 2010, p17), a concept first introduced by Hockett (1960) provides a promising tool to do this.

## **1.2 The multi-component approach to language evolution**

The multi-component approach to language draws on Darwin's concept of descent with modification and uses a broad comparative method to identify shared mechanisms underlying human and animal behaviour. There are three core concepts within this approach. The first is that the language faculty is a complex system made up of many separate subcomponents, and that each subcomponent may have a different function (not necessarily limited to

communication) and a different evolutionary pathway (Hauser et al. 2002). The second is that the evolving language faculty would have been more likely to co-opt pre-linguistic subcomponents than to evolve entirely novel language-specific cognitive modules, and that these pre-linguistic mechanisms can be thought of as pre-adaptations to language (Hurford 2003). The third is that while language is unique in being composed of all of these subcomponents, some (possibly all) of these single subcomponents or pre-adaptations may be present in other taxa (Zuberbühler 2003; Schoenemann 2005). Comparative studies of animal communication and cognition can therefore be used to identify whether subcomponents of language, or biological pre-adaptations to these subcomponents, are shared with other extant species (Hauser and Fitch 2003; Weiss and Newport 2006; Fischer 2010), offering insights into the cognitive tool-kit with which our hominin ancestors were likely armed prior to the emergence of language.

In order to use the comparative approach to identify whether subcomponents of language are shared with other species, it is necessary to distinguish what at least some of these subcomponents might be, or as Chomsky puts it, "there is little point in speculating about the process of acquisition without a much better understanding of what is acquired." (Chomsky 1959, p55). Hockett (1960) proposed an insightful early list of unique and shared "design features" of human speech (he focused on language as expressed in the vocal domain); but more than 50 years later, there is still no definitive list of what constitutes the language faculty (Hauser et al. 2002). Nevertheless, there is a general agreement that modern language requires vocal flexibility (Hammerschmidt and Fischer 2008; Oller and Griebel 2008), and a shared representational state between the speaker and listener (Rendall et al. 2009; Fischer and Hammerschmidt 2011) based upon semantics (meaningful words and sentences) and pragmatics (the speaker's intended and/or the listener's inferred meaning; Fitch 2010).

The aims and objectives of this PhD thesis are based upon the theoretical concepts and current understanding of vocal flexibility, semantics and pragmatics in humans and animals, and I will introduce each of these topics in turn in the sections below. For the most part I will concentrate on the comparative perspective offered by studies of our nearest relatives, the primates, although I will incorporate findings from a wider range of species when this is important to the point being made.

### 1.2.1 Vocal flexibility

*It is a very inconvenient habit of kittens that, whatever you say to them, they always purr. If they would only purr for "yes" and mew for "no", or any rule of that sort, so that one could keep up with a conversation! But how can you talk with a person if they always say the same thing? (Lewis Carroll, Alice through the looking glass)*

An integral feature of human speech is its intrinsic flexibility, seen in its open-ended creativity and the degree to which it varies between populations, with regional differences in language type, dialect and accent (Lameira et al. 2010). In the European Middle Ages the standard account for this diversity of spoken languages was based upon the biblical story of the Tower of Babel wherein God is responsible for confounding the one language of all Earth; "let us go down, and there confound their language, that they may not understand one another's speech" (Genesis 11:5-6). A more scientific approach has led to the proposal that the emergence of communicative flexibility was one critical step in the evolutionary course towards modern language (Oller and Griebel 2008). Flexibility in speech can be attributed to the ability to modify vocal structure as a result of learning (vocal learning), and also to the ability to produce and respond to words in novel contexts (contextual learning; Janik and Slater 2000). Following the scientific approach, numerous comparative studies have thus been carried out to identify to what extent other species demonstrate flexibility in the structure and use of their calls.

#### Call production

Vocal learning encompasses learnt changes in call structure, and in humans is apparent in two forms: learned acquisition when development of the species-typical vocal repertoire is dependent on auditory experience through a process of vocal imitation, and social modification when existing vocalisations are modified as a result of auditory experience (Boughman and Moss 2002). The learnt acquisition of novel sounds has been identified in only a few distantly related taxonomic groups, including songbirds (reviewed in Doupe and Kuhl 1999; Wilbrecht and Nottebohm 2003), marine mammals (reviewed in Nottebohm 1972; Janik and Slater 1997) and elephants (*Loxodonta africana*, Poole et al. 2005; *Elephas maximus*, Stoeger et al. 2012). Non-human primates are notably absent from this group (Egnor and Hauser 2004), as they appear to acquire species-typical vocalisations even when deprived of normal auditory experience by social isolation (Winter et al. 1973), deafness (Hammerschmidt et al. 2000; 2001), or cross-fostering (Owren et al. 1992). In humans, the ability to exercise voluntary control over the spectral patterning of words and the production of novel sounds depends on a direct connection between the primary motor cortex and the nucleus ambiguus which in turn controls the

laryngeal motorneurons (Kuypers 1958); a similar direct connection is found between forebrain motor areas and neurones which control syringeal movements in songbirds (Wild 1993). This connection is missing in primate vocal production (Jürgens 1976a), a difference that likely accounts for their inability to produce calls outside of the species-typical vocal repertoire (Jürgens 2009). Interestingly, this direct connection is also not involved in the production of human non-verbal sounds, such as laughs, cries and shrieks; thus at the neurological level, animal calls appear more similar to this group of innate vocalisations than to words (Hage 2010).

In spite of this missing connection, analyses of fine-scale acoustic structure find that primate vocal structure may undergo social modification, a process described by Seyfarth and Cheney as "modification within constraints" (1997). Some ontogenetic vocal changes are likely to be the result of maturational development (Lieblich et al. 1980; Hammerschmidt et al. 2000), but captive studies indicate that at least some primate species exercise limited control over call duration and amplitude (Sutton et al. 1973) and there is also growing support that convergence or divergence of fine-scale temporal and spectral call structure can arise as a result of social modification in pygmy marmosets (*Cebuella pygmaea*, Elowson and Snowdon 1994; Snowdon and Elowson 1999), Barbary macaques (*Macaca sylvanus*, Fischer et al. 1998) and chimpanzees (*Pan troglodytes*, Marshall et al. 1999; Crockford et al. 2004). Such convergence of call structure may constitute a form of vocal accommodation whereby vocalisers modify the temporal and/or spectral characteristics of their vocal output to resemble that of social partners, thus exhibiting parallels to speech accommodation found in humans (described in Fischer 2003). Subtle changes in the temporal characteristics and amplitude of calls can be attributed to the modulating activity of the periaqueductal gray (PAG, located in the midbrain) which acts as a primary vocalisation-eliciting area in all terrestrial mammals studied so far (Jürgens 1998). Within this process, the PAG receives input from various limbic pathways and controls the initiation and amplitude of vocal output on the basis of this incoming stimulation; strikingly similar neurological pathways are responsible for the production of human non-verbal sounds (Jürgens 1976b; Dujardin and Jürgens 2006; Jürgens 2009). The neural mechanisms underlying the subtle spectral differences observed in animal calls are, however, less well understood (Hammerschmidt and Fischer 2008), but may be related to sensory-motor integration or action based learning (Fischer 2008; Fischer 2010).

### **Call use**

Non-human primates are commonly described as having more control over call use than call structure (Seyfarth and Cheney 2010), in particular the ability to control whether or not a call is produced. It has been found in a number of species that the presence or identity of other

individuals affects the incidence of call production (alarm calls: Gyger et al. 1986; le Roux et al. 2008; food calls: Evans and Marler 1994; Di Bitetti 2005), a phenomenon known as an "audience effect" which appears to be a common trend in animal communication in the wild (Zuberbühler 2008; Seyfarth and Cheney 2010). Vervet monkeys (*Chlorocebus pygerythrus*) provide a good example, as it has been observed that they are less likely to produce alarm calls when no conspecifics are present (Cheney and Seyfarth 1985a), and that females are more likely to produce alarm calls when their own offspring are present than when accompanied by unrelated juveniles (Cheney and Seyfarth 1992). Operant conditioning tasks carried out under captive settings strengthen claims of volitional control over vocal production, by showing that primates are able to initiate or inhibit calling in response to an external cue (Leander et al. 1972; Aitken and Wilson 1979; Koda et al. 2007), and to alter call timing to avoid overlapping with conspecific calls (Hage 2013) or background noise (Egnor et al. 2007). Looking again at the neural mechanisms underlying call production in primates, the anterior cingulate cortex can be thought of as being one level up from the PAG and as playing a role in enabling the volitional control of internally triggered vocalisations (Sutton et al. 1974; Jürgens 1992). Interestingly, background noise also causes individuals to alter their rate of calling, as well as call duration and amplitude (Brumm et al. 2004; Egnor and Hauser 2006) indicating that the influential role of the PAG in controlling all of these features may sometimes blur the line between flexibility in call structure and call use (Fischer 2003). It is unclear to what degree this control is independent from motivational triggers; humans' ability to stifle a laugh clearly demonstrates the important point that innate and emotional calls can also demonstrate a degree of voluntary control (Fitch 2010).

Limited flexibility is also seen in the context in which a call is produced, for example the development of "correct" alarm call use by wild immature vervet monkeys (Seyfarth and Cheney 1986). Importantly, however, the range of contexts in which animal calls are naturally produced is generally quite restricted (Seyfarth and Cheney 1997), which suggests that rather than individuals learning which contexts calls should be produced in, vocal development is related to processes of stimulus recognition (Wheeler and Fischer 2012) and/or habituation (Owren et al. 2011), brought about by learnt associations/disassociations between the event eliciting calling and an experienced emotional state. Thus while such flexibility in call production and call use may represent a pre-cursor to the flexibility observed in human speech, it may also be more akin to the limited flexibility observed in the production of innate non-verbal sounds (Slocombe and Zuberbühler 2007). Distinguishing between these alternative explanations is highly relevant to discussions of animal semantics.

### 1.2.2 Semantics

*"It's a stupid enough name! Humpty Dumpty interrupted impatiently. What does it mean? MUST a name mean something? Alice asked doubtfully. Of course it must, Humpty Dumpty said with a short laugh: My name means the shape that I am-and a good handsome shape it is, too. With a name like yours, you might be any shape, almost." (Lewis Carroll, Alice in Wonderland)*

Words have meaning in that they represent something other than themselves and semantics encompasses the study of this meaning (Deacon 1997; Hurford 2007; Fitch 2010). More specifically, words have symbolic meaning because the relationship between the word and that to which it refers is arbitrary and based on a set of conventional rules (Peirce 1958). The use of symbols has been put forward as a fundamental step in the evolution towards modern human language (Deacon 1997; Jackendoff 1999; Christiansen and Kirby 2003), so that following on from concepts of vocal flexibility, a second central question in animal communication research has been whether vocal signals are mainly a component of the signaller's emotional state or whether they refer arbitrarily to specific objects or events in the environment (Marler 1977).

#### **Emotional calls and functionally referential signals**

In lieu of Darwin's influential book "The expression of the emotions in man and animals" (Darwin 1872), animal communication was traditionally considered to encode cues relating to the caller's emotional state (Rowell and Hinde 1962; Lancaster 1975; Premack 1975). This perspective implied that "variation in the sender's internal state will be reflected by continuous gradation in the physical properties of the signal produced. Signals will be evoked under a very wide range of environmental circumstances and will consequently only be interpretable with the aid of contextual information" (Evans 1997, p. 4). Importantly an individual's internal state can usefully be broken down into smaller constituent parts; for example Todt (1986; described in Fischer 2011) recognised the internal state as consisting of an affective component related to the individual's evaluation of the environment, a motivational component related to the individual's action tendencies, and an arousal component related to the individual's likelihood and urgency to respond. Appraisal theories of emotion (Arnold 1960; Lazarus 1966) take a similar approach. Within the framework of appraisal theories, an individual's emotional state arises from an appraisal of the environment, and emotional state can be broken down into four components; the individual's motivational state (readiness to act), peripheral physiology (e.g. hormonal and nervous system changes, likely related to caller arousal), expressive behaviour (e.g. vocalisations), and feeling (subjective experience), for more detail see Moors and

colleagues (2013). Both approaches are therefore useful for assessing how different components of an individual's internal/emotional state may affect vocal output, and potentially also for identifying parallels between the vocal expression of emotions in animals and humans (Fischer 2011); points I return to within my general discussion. Within this thesis I use the terminology of appraisal theory, although because I am concerned with emotional states in animals rather than humans, my use of the term emotion does not imply feelings in the sense of subjective and conscious experiences, but rather in the sense of basic emotional urges (Panksepp 2011).

The finding that vervet monkeys' predator-specific alarm calls elicit predator-appropriate responses in conspecifics, even in the absence of contextual cues (Seyfarth et al. 1980a; Seyfarth et al. 1980b) was the first suggestion that animal repertoires might contain symbolic as well as emotion-based calls. But there are important limitations to how much can be understood about a call's semantic value based only on the responses it evokes in others (Seyfarth and Cheney 2003a; Fischer 2010; Wheeler et al. 2011); in acknowledgement of this, the term "functionally referential" was subsequently introduced (Macedonia and Evans 1993). The framework of functional reference provides criteria to determine where to place animal vocalisations along a continuum from emotion-based calls reflecting the signaller's motivational state to functionally referential calls indicative of an external object or event. More recently, the concept of a single continuum has been described as a false dichotomy (Seyfarth and Cheney 2003b; Rendall et al. 2009; Fischer 2011), as an emotion-based system of call production may still function referentially if the signal receiver is able to associate affect-based cues with an external object or event (Premack 1975). To identify to what extent animal signals might constitute a semantic precursor to words, it is therefore necessary to determine the relation between animal calls and what they denote from the perspective of the signaller and the receiver (Smith 1977). Below I do this using the criteria laid out for functionally referential signals.

### **Call production**

Production criteria for functional reference have historically been that calls produced in different contexts should exhibit discrete rather than graded differences in call structure (Marler et al. 1992; Evans and Marler 1995), and that calls should demonstrate stimulus specificity in that eliciting stimuli should belong to a cohesive category (Macedonia and Evans 1993). Many animal vocalisations (and also human laughs and cries) do not meet these production criteria and in these cases there is a general consensus that vocal production is strongly related to the signaller's emotional state (Marler et al. 1992).

Graded variation in acoustic parameters can be brought about by differences in caller arousal via involuntary changes in phonation, resonance, articulation, respiration and salivation

(Briefer 2012). Based on an understanding of the underlying physiology of sound production, predicted changes are that as caller arousal increases, call duration gets longer, calls are produced more rapidly, calls become less tonal, and call frequencies get higher and more variable (Scherer et al. 2003; Briefer 2012). Correlations such as these have been found in redfronted lemurs (*Eulemur fulvus*; Fichtel and Hammerschmidt 2002), squirrel monkeys (*Saimiri sciureus*; Fichtel et al. 2001), common marmosets (*Callithrix jacchus*; Yamaguchi et al. 2010), Barbary macaques (Fischer et al. 1995), chacma baboons (*Papio ursinus*; Meise et al. 2011) and chimpanzees (Slocombe and Zuberbühler 2007). That similar results are also found in studies of human vocalisations (reviewed in Scherer 1989) and in a number of species outside of the primate taxa (*L. africana*, Soltis et al. 2005; *Crocuta crocuta*, Theis et al. 2007; *Canis familiaris*, Taylor et al. 2009) indicates that comparable effects of caller arousal on call structure are widespread at least in terrestrial mammals that share a similar vocal production anatomy (Fitch 2003). In some species, caller arousal has also been proposed to influence the type of call produced; such calls demonstrate low stimulus specificity because they are related to different levels of risk (often predator distance) rather than a specific external stimulus, and within the alarm calling context this is referred to as a response-urgency system (Owings and Hennessy 1984). A correlation between risk and the type of call produced has been found in rodents and some other species (*U. beldingi*, Robinson 1980; *Marmota caudata*, Blumstein 1995; *Marmota flaviventris*, Blumstein and Armitage 1997; *Acrocephalus scirpaceus*, Welbergen and Davies 2008; *Mungos mungo*, Furrer and Manser 2009). In primates, predator distance seems more likely to affect subtle structural differences (*M. sylvanus*, Fischer et al. 1995; *Cercopithecus diana*, Zuberbühler 2000a) or call rate (*Cebus apella nigrurus* Wheeler 2010) than the type of call produced, although the degree of risk (inferred from predator distance) does appear to affect the frequency at which different call types are produced in at least one primate species (*Cercopithecus mitis*, Murphy et al. 2013).

In contrast to calls produced in response to the degree of risk experienced, the calls of many animal species instead demonstrate predator-specificity (*Lemur catta*, Macedonia 1990; *Gallus gallus*, Evans et al. 1993; *Cercopithecus campbelli*, Zuberbühler 2001; *Suricata suricatta*, Manser et al. 2002; *Cynomys gunnisoni*, Slobodchikoff 2002; *Dendroica petechia*, Gill and Sealy 2004; *Saguinus fuscicollis* and *S. mystax*, Kirchhof and Hammerschmidt 2006). To a lesser extent stimulus-specific calls have also been identified within contexts of food discovery (reviewed in Clay et al. 2012) and during social interactions (Gouzoules et al. 1984; Semple et al. 2002; Faragó et al. 2010; Struhsaker 2010). It has been claimed that these calls are "more than mere expressions of emotional state" (Hauser 1997, p. 509) and that such signals may have been a precursor to semanticity and human words (Zuberbühler 2003; Fedurek and Slocombe 2011;

Townsend and Manser 2012). But to what extent the production of such calls can be attributed to mechanisms unrelated to the caller's emotional state when parameters such as motivational state are considered in addition to caller arousal remains far from clear (Seyfarth and Cheney 2003a; Fischer 2011). To shed light on this important question, more studies are needed to look not at how distinguishable a call is (a question far more relevant to how calls are perceived) but at how call structure varies within and between call types not limited to a single context.

### **Call perception**

The criterion given for functional reference on the receiver's side is that calls should demonstrate contextual independence; that is, the signal should elicit an appropriate response even in the absence of supporting contextual cues (Macedonia and Evans 1993). Subsequent to Seyfarth and colleagues' descriptions of predator-appropriate responses to alarm calls in vervets (1980a; 1980b), numerous other species have been shown to respond with predator-specific behaviour to alarm calls (*Cercopithecus diana*, Zuberbühler 2000a; *Suricata suricata*, Manser et al. 2001; *E. rufus* and *Propithecus verreauxi*, Fichtel and Kappeler 2002; *Saguinus fuscicollis* and *S. mystax*, Kirchhof and Hammerschmidt 2006; *Cynomys gunnisoni*, Kiriazis and Slobodchikoff 2006), and there is some support that food-associated calls elicit enhanced foraging behaviour (Slocombe and Zuberbühler 2005; Kitzmann and Caine 2009).

With regards to the cognitive implications of these findings, one suggestion is that calls elicit the responses they do as a result of a direct effect of the sound on the receiver (Rendall et al. 2009). Whilst the structure of calls does play a role in how receivers respond (a topic discussed earlier in relation to the form and function of signal structure), playback experiments indicate that many receiver responses cannot be attributed solely to an unconditioned reaction to the acoustic properties of a call; rather they require a degree of learning (Hammerschmidt and Fischer 1998a; Fischer et al. 2000; Zuberbühler 2000b; Semple 2001). To what degree signal receivers' responses are the result of learning to associate particular calls with an external referent, versus affect-conditioning (a learnt association between an event and an experienced emotional state, Owren and Rendall 1997) is, however, often less clear.

A point of interest related to the learning of functional categories is how within- and between-category variation is perceived, especially in the case of graded signalling systems. The ability to discard within-category variation and distinguish between-category variation has been labelled "categorical perception" (Harnad 1987), a phenomenon which, when applied to the perception of sound, was first thought to be unique to speech processing (Liberman 1957). Subsequently it has been found that some animals are also capable of categorical perception of human and conspecific vocalisations (reviewed in Fischer 2006).

### 1.2.3 Pragmatics

*“When I use a word, Humpty Dumpty said in a rather scornful tone, it means just what I choose it to mean—neither more nor less. The question is, said Alice, whether you CAN make words mean so many different things”. (Lewis Carroll, Alice in Wonderland)*

Linguistic meaning is described in Grice's (1957) seminal paper as depending not only on the relationship that exists between a word and that which it represents (the key concept of semantics), but also on the fact that both the signaller and the receiver take each another's state of mind into account when communicating. For instance, if a man tells his wife "you look nice", the meaning of the sentence is not so much that the woman is looking nice, but rather that her husband thinks that she looks nice and that he intends that she should know this. Linguistic pragmatics can be described as the study of such intended and/or inferred meaning (Fitch 2010). A second form of inference also exists within human communication, and I will refer to this form as "contextual pragmatics" to highlight that in this case receivers do not attribute word meaning from an inferred intention of the signaller, but rather that they interpret an utterance in relation to the context within which it was made (Wheeler et al. 2011; Semple and Higham 2013). An example of such inference is the utterance "the chicken is ready to eat"; whereby the inferred meaning of this phrase would likely be different if the speaker was walking into the garden with a bag of chicken feed in their hand than if they were speaking whilst taking a roast chicken out of the oven.

Importantly, linguistic and contextual pragmatics require different levels of intentionality. Viewed within Dennett's (1983) framework of "intentional system theory", linguistic pragmatics displays second-order (or higher) intentionality in that it requires that the signaller and the receiver understand something of their own and/or another's mental state ("theory of mind"; Premack and Woodruff 1978), whereas contextual pragmatics requires only first-order intentionality, the possession of a mental state (belief, desire, or goal). Comparative studies can be used to assess whether animals are capable of linguistic and/or contextual pragmatics to better understand the emergence of this component of language.

#### **Linguistic pragmatics**

With regards to linguistic pragmatics, it is useful, as with semantics, to assess empirical studies of call production and call perception separately. Thus relevant questions within animal communication are firstly, do signallers produce vocal signals with the intent of altering another's state of mind, and secondly do receivers recognise intentions in others. From the perspective of call production, animals do not appear to take other's knowledge state into

account when calling (Cheney and Seyfarth 2005; Fischer 2008). Support for this conclusion comes from studies of captive female macaques (*Macaca fuscata* and *M.mulatta*) with infants who were presented with a "predator" (in fact, Marc Hauser disguised as a veterinarian) or with hidden food; importantly, experiments varied in whether the female's infant was or was not aware of the predator or food, and showed that the infant's knowledge state had no effect on the mother's calling behaviour (Cheney and Seyfarth 1990). Similarly, chacma baboons appear to produce contact barks in response to conspecific's contact barks only when they themselves are separated from the group (Cheney et al. 1996; Rendall et al. 2000), further supporting the hypothesis that primates, and most likely other animals, do not deliberately vocalise in order to inform others (Cheney and Seyfarth 1996). Turning to signal receivers, I have found no empirical study that has looked at whether signal receivers infer intention on the part of the signaller. There have, however, been numerous studies carried out within the field of social cognition to assess whether animals understand anything about the unobservable mental states of others. Though still a contentious issue, the general consensus is that animals lack a full blown theory of mind (Penn and Povinelli 2007), identifying this as an important component of the language faculty that most likely arose after hominins split from other apes (Cheney and Seyfarth 1998; Cheney and Seyfarth 2005).

### **Contextual pragmatics**

In the above paragraph I described how the attribution of call meaning in animals does not appear to be influenced by any concept of signaller intentions. In contrast, there is growing support that across a wide range of species, signal receivers make pragmatic inferences on the basis of both the external setting at the time of signal production (I will refer to this as "public contextual information") and the receiver's previous experience (I will refer to this as "private contextual information"). One example of how public contextual information can be incorporated into receiver responses comes from a study of chacma baboons that used playback experiments to show that, while conspecific responses to move-grunts and infant-grunts were influenced by call structure (i.e. move grunts tended to elicit typical move behaviour while infant grunts elicited infant-handling behaviour), calls played back in an "incorrect" context reduced the likelihood of call type-consistent behaviours (Rendall et al. 1999). Such effects have also been found in alarm calling contexts, whereby tufted capuchins (*Cebus apella nigrinus*) respond less strongly to alarm calls produced in a feeding than non-feeding context, possibly as a counter-strategy to functionally deceptive alarm calling (Wheeler and Hammerschmidt 2013). Examples of private contextual information use are perhaps even broader, having been identified in the foraging behaviour of ants (*Lasius niger*, Grüter et al. 2011), the territorial

behaviour of male song sparrows (*Melodia melospiza*, Akçay et al. 2009), and the social interactions of female chacma baboons (Engh et al. 2006). Viewed collectively, studies such as these suggest that social intelligence and the ability to make simple pragmatic inferences may constitute an evolutionary precursor to the emergence of linguistic pragmatics (Cheney and Seyfarth 2005). Less well understood, however, is how call structure and call context interact to influence the attribution of meaning. This thesis aims to shed light on such questions through the study of vocal communication, and particularly alarm calling, in African green monkeys (*Chlorocebus*).

### **1.3 African green monkeys: A model taxon**

There is ongoing debate within the scientific literature concerning the taxonomy and nomenclature of African green monkeys. It is therefore important to start by clarifying that, whilst many previous and some current studies consider African green monkeys to be a single species *Cercopithecus aethiops*, in this thesis I follow the taxonomy of Groves (2001; 2005). Here, African green monkeys are assigned to their own genus (*Chlorocebus*) and the phylogenetic species concept is applied to recognise six species, *C. aethiops* (grivet), *C. djamdjamentis* (Bale monkey), *C. sabaues* (green monkey), *C. cynosuroides* (malbrouck monkey), *C. tantalus* (tantalus monkey), and *C. pygerythrus* (vervet).

The genus of African green monkeys constitutes a superb model taxon for comparative studies aiming to identify socio-cognitive precursors to language. They have historically played an influential role in comparative studies of non-human vocal communication; in particular they are subject to high predation pressure (Cheney and Seyfarth 1992) and numerous studies have been carried out describing the production and perception of alarm calls (Seyfarth et al. 1980a; Seyfarth and Cheney 1980; Owren and Bernacki 1988; Owren 1990a; Seyfarth and Cheney 1990; Hauser 1993b). Non-alarm calls have also been studied (Struhsaker 1967b; Cheney and Seyfarth 1980; Cheney and Seyfarth 1982b), and collectively these present a solid background of knowledge upon which to base further research. Furthermore, like many anthropoid primates, African green monkeys form stable groups within which individuals form enduring social relationships (Struhsaker 1967a; Seyfarth 1980; Fairbanks 1980; Fairbanks and McGuire 1985); being semi-terrestrial and easily habituated they are also easy to observe and extensive observational and experimental studies have demonstrated considerable social knowledge (Cheney and Seyfarth 1985b; Cheney and Seyfarth 1986) and social learning capabilities (van de Waal and Bshary 2010; van de Waal et al. 2010; van de Waal et al. 2013).

Regarding vocal flexibility, i.e. the amount of control individuals exercise over their vocal production, African green monkeys are widespread over sub-Saharan Africa (Lernould 1988; Kingdon 1997), and exhibit extensive morphological differences (Dandelot 1959; Napier 1981) and social flexibility at the individual level (Henzi et al. 2013). Thus there is potential for species and subspecies to vary in their vocal output as a result of directional genetic divergence or drift and/or social modification. Nevertheless, the vast majority of empirical data concerning vocal communication in this genus comes from a single population of vervet monkeys in East Africa (reviewed in Cheney and Seyfarth 1992), and inter-specific comparison of vocal structure is limited to qualitative descriptions (Struhsaker 1970). It is thus unclear to what degree acoustic variation across the genus relates to phylogenetic relatedness.

In terms of call meaning, the alarm calling system of the vervet monkey constitutes the classic example of functional reference in a non-human animal, with the production of acoustically distinct "leopard", "snake", and "eagle" alarm calls eliciting predator-appropriate responses in conspecifics even in the absence of supporting contextual cues (Seyfarth et al. 1980a). Additional studies in captivity have found that structural differences exist between female alarm calls produced to snakes and eagles (Owren and Bernacki 1988) and that vervets can discriminate between natural and synthetic exemplars of these calls (Owren 1990a; Owren 1990b). On the basis of these findings it has been surmised that vervet monkeys produce calls in a predator-specific fashion, that they can distinguish these calls by ear, and that they have learnt to associate them with different categories of predator. Nonetheless, many questions remain regarding the cognitive mechanisms underlying these behaviours.

### **1.4 Study aims and approaches**

So far in the thesis I have introduced the alarm calling system of the vervet monkey as the classic example of functional reference in animals and described the important role it has played in the search for the origins of language. I have also introduced findings that indicate that vocal production in primates is largely innate and shaped by the caller's emotional state, while call perception is influenced by learning. In light of these findings, the overall aim of this thesis is therefore to increase understanding of the mechanisms underlying the production and perception of functionally referential alarm calls in African green monkeys in order to assess the relevance of functional reference to the emergence of vocal flexibility and meaning in language. Three inter-related studies were carried out to achieve this.

In study 1 (**Chapter 2**) I investigate vocal flexibility within the African green monkey genus to gain insight into the degree of control individuals exercise over call production. To this end I examined geographic variation in the adult male alarm bark, providing the first quantitative acoustical analysis of vocal differences within and between African green monkey species. Furthermore I assessed male responses to conspecific and heterospecific (but congeneric) alarm barks to examine the function of acoustic differences in this call.

In study 2 (**Chapter 3**) I focus on alarm call production in the vervet monkey to better understand the parallels existing between the production of these alarm calls and human words. To achieve this I analysed original vocal data sets and additional more recent recordings to quantitatively evaluate the extent of structural gradation between alarm calls produced in response to different predator types, and the production specificity of alarm calls when compared first with other alarm calls, and second with calls produced in non-predator contexts.

In study 3 (**Chapter 4**) I explore call perception in the green monkey to increase knowledge of meaning attribution and pragmatic inference within the natural alarm calling system of a wild primate. With this aim, I used field experiments to systematically control the effects of call context and call structure so as to evaluate the role of each in determining green monkey responses to a graded alarm call.

In a general discussion (**Chapter 5**) I summarise the major results of these studies and compare them to what is known from other studies. I then discuss the relevance of my findings in relation to questions of language evolution, and in a final section, provide a short conclusion and outlook for future research.



## Chapter 2

# Limited geographic variation in the acoustic structure of and responses to adult male alarm barks of African green monkeys

Tabitha Price<sup>1,2,3\*</sup>, Oumar Ndiaye<sup>4</sup>, Kurt Hammerschmidt<sup>1</sup>, Julia Fischer<sup>1,2</sup>

<sup>1</sup>Cognitive Ethology Lab, German Primate Center, Germany

<sup>2</sup>Courant Research Centre for the Evolution of Social Behaviour, Georg August University of  
Göttingen, Germany

<sup>3</sup>Applied Behavioural Ecology and Ecosystem Research Unit, UNISA, South Africa

<sup>4</sup>Direction de Parc National de Niokolo-Koba, Sénégal

Behavioral Ecology and Sociobiology (Under Review)

### **Abstract**

The diversity of human languages across the globe is a remarkable feature of our species, resulting from our capacity for rapid vocal learning and cultural evolution. Identifying geographic variation in non-human primate vocalisations and understanding the underlying causal mechanisms are thus important steps to help uncover evolutionary precursors to language. This is specifically true for alarm calls, which play an important role in the language origin debate. This study set out to investigate variation in the alarm bark of the widely distributed African green monkey (*Chlorocebus*) and to assess how this variation is perceived. We first compared the adult male barks of green monkeys (*C. sabaesus*) and two subspecies of vervet (*C. pygerythrus pygerythrus* and *C. pygerythrus hilgerti*). We then carried out playback experiments to test whether adult male vervets would distinguish between the barks of an own group, unknown conspecific, and green monkey male. Whilst similar in overall structure, the barks of African green monkeys could be distinguished from vervet barks with a high degree of accuracy; the barks of vervet subspecies could also be discriminated, but to a lesser degree. Males responded most strongly to barks from unknown conspecific males, and exhibited responses typical of leopard-avoidance and male-male aggression. Taken together, these findings indicate that variation in alarm calls can be best explained by phylogenetic distance, and that intra- and inter-species differences are socially relevant. Moreover, barks may function as an alarm and display call, explaining the previously observed strong sexual dimorphism in alarm calls in this genus.

## Introduction

Human speech displays extensive regional differences in language type, dialect and accent, and it has been proposed that the emergence of basic flexibility in signal structure was a critical step in the evolution of modern language (Oller and Griebel 2008). Vocal communication in non-human animals also exhibits geographic variation (Weilgart and Whitehead 1997; Slobodchikoff et al. 1998; Bradbury et al. 2001; Davidson and Wilkinson 2002; Smith and Hunter 2005; Delgado 2007), but whilst speech patterns are strongly influenced by learning, the ability to produce novel vocalisations as a result of experience has been identified in only a few distantly related taxa (Janik and Slater 1997). In non-human primates (hereafter primates) this ability is notably absent (Egnor and Hauser 2004).

Between-species variation in primate loud-call structure is generally attributed to genetic differences (Oates and Trocco 1983; Brockelman and Schilling 1984; Méndez-Cárdenas et al. 2008; Wich et al. 2008; Thinh et al. 2011; Meyer et al. 2012), and species-specific differences in such calls can be used as a non-invasive tool for discriminating between cryptic species (Nietsch and Kopp 1998; Zimmermann et al. 2000). Conversely, it has been suggested that alarm calls should be under stronger stabilising constraints (Struhsaker 1970); but whilst this prediction is supported by a lack of variation between the alarm calls of two mouse lemur species (Zimmermann et al. 2000), rodent alarm calls appear to be less constrained (Slobodchikoff and Coast 1980; Conner 1982; Eiler and Banack 2004; Schneiderová and Policht 2011). Quantifying the geographic variation present in alarm calls within and between other closely related primate species, and investigating how vocal variation affects receiver responses will thus offer a clearer understanding of the dynamics of primate vocal evolution.

African green monkeys (*Chlorocebus*) provide an excellent model to study such processes. This group of monkeys is one of the most widespread African primates, distributed over much of sub-Saharan Africa (Lernould 1988; see Figure 2.1). Extensive morphological differences have been recorded across this range (Dandelot 1959; Hill 1966; Napier 1981) and these differences have been used to split the genus into four monotypic, and two polytypic species (Groves 2001; 2005). Whilst taxonomy within the genus is still disputed (Grubb et al. 2003), recent analyses of mtDNA diversity clearly separate the green monkey (*C. sabaesus*) in the West from all other species (Haus et al. 2013), and propose that the initial split within African green monkeys occurred between this Western clade and all other lineages approximately 2.81-2.76 MYA (Wertheim and Worobey 2007). Analyses of mtDNA also support genetic separation within the polytypic vervet (*C. pygerythrus*) taxon, between mainland subspecies *C. p. hilgerti* ranging from Ethiopia to northern Tanzania, and *C. p. pygerythrus* in Southern Africa (Haus et al. 2013). This

variation can be attributed to a more recent period of rapid diversifications within the genus, occurring approximately 1.59-1.48MYA (Wertheim and Worobey 2007).

Despite these morphological and genetic differences, vocal comparisons amongst African green monkey populations were, until now, limited to qualitative descriptions. These observations suggested that vocalisations are spectrally stable among vervet subspecies and between vervet and tantalus (*C. tantalus*) populations, but that they may exhibit differences in temporal characteristics (Struhsaker 1970). A quantitative analysis of intra- and inter-species vocal differences and how African green monkeys respond to the vocalisations of other populations was, however, lacking. This is somewhat surprising, as the alarm calls of East African vervets constitute one of the most well-known examples of functionally referential signals (Seyfarth et al. 1980a), and studies of geographic variation in this genus may offer insights into the flexibility present in their vocal behaviour, and thus shed light on the mechanisms that give rise to context-specific calling.

In this study we carried out structural analyses of bark calls produced by adult male South African vervets (*C. p. pygerythrus*), adult male East African vervets (*C. p. hilgerti*), and adult male green monkeys (*C. sabaesus*) to investigate variation between and within species, including the variation between individuals. Following this, we tested how South African adult male vervets perceive the barks of own-group males and unknown males of the same subspecies, and unknown male green monkeys. The adult males of all three populations have been reported to produce barks in response to territorial predators, and playback experiments have shown that East African vervets frequently respond to conspecifics' alarm barks by climbing a tree, an appropriate leopard-avoidance behaviour (Seyfarth et al. 1980a). Similar leopard-avoidance behaviour was observed in green monkeys, with subjects always climbing more than 2m into a tree in responses to a leopard model (Price and Fischer 2013). Thus if South African vervets recognise intra- and inter-specific bark calls as indicating the presence of a terrestrial predator, they should respond by climbing to a height of more than 2m into a tree. Barks may also be given during aggressive encounters (Struhsaker 1967b; Galat and Galat-Luong 1976; Cheney and Seyfarth 1992), which frequently entail males chasing after another male (TP personal observation). Nevertheless, acoustic analysis of South African vervet barks suggest that whilst calls produced in response to leopards and during aggression are acoustically similar, they are in most cases distinguishable (Price et al. under review).

The aim of this study was thus to identify the degree of flexibility present within the structure of bark calls, to investigate how this flexibility is perceived, and to offer insights into the function of adult male barks. Given the previous observations of Struhsaker (1970) and that the structure of primate alarm calls may be particularly constrained (Zimmermann et al. 2000),

we predicted to find little variation between and no variation within species. Considering the presumed stabilising selection, and the costs of not responding to a putative alarm call, we expected calls to cause males to climb into trees, and furthermore we expected relatively little variation in listeners' responses to the playback of calls with different origins.

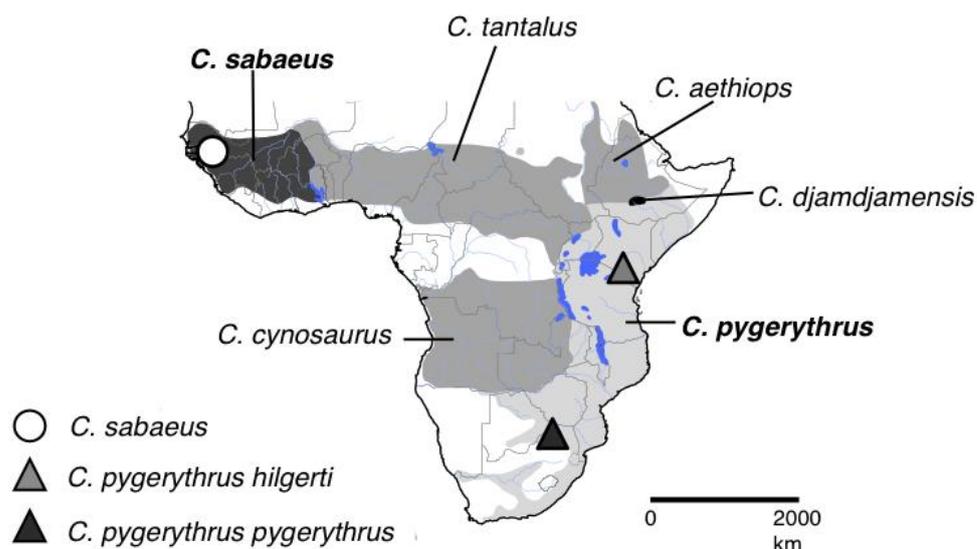
## **A. Geographic and individual differences in call structure**

### **Data collection**

Adult male bark vocalisations were recorded from green monkeys, East African vervets, and South African vervets at three geographically distant study sites within the range of African green monkeys (Figure 2.1). All study subjects were habituated to human presence and were recognised individually from natural markings on the face and body. The barks of East African vervets were recorded by Thomas Struhsaker (June 1963- May 1964), and Robert Seyfarth and Dorothy Cheney (1977-1988) as part of their earlier studies on several free ranging groups within the semi-arid acacia savannah of Amboseli National Park (2°39'49S; 37°15'16E) in Kenya, and these calls were subsequently made available for inclusion within this study. Green monkey barks were recorded by TP over two field seasons (January-June 2010 and 2011) from four free-ranging groups and two solitary males within Niokolo Koba National Park (13°01'34"N, 13°17'41"W), an area in southeastern Senegal consisting mainly of Sudano-Guinnean savannah interspersed with woodland and gallery forest (Frederiksen and Lawesson 1992). South African vervet barks were recorded by TP and ON (January-June 2012) from five free-ranging groups located within the Loskop Dam Nature Reserve (25°25'18S; 29°18'29E) in South Africa, which contains a mixture of open grassland, acacia dominated woodland and low mountains with open tree savannah (Filmlalter 2010).

In all studies, adult male barks were recorded ad libitum when the context of calling could be confirmed as the presence of a feline terrestrial predator, either by observing the predator or hearing its vocalisations. Whilst the natural occurrence of bark calls was not uncommon, it was often not possible to confirm whether a terrestrial predator was present at these times. Following numerous studies that have successfully used the presentation of predator models to elicit alarm calling (e.g. Coss et al. 2007; Arnold et al. 2008; Wheeler 2010), spontaneous barks from all field sites were supplemented with barks produced in response to leopard models. Altogether, five different leopard models were used. All East African vervet calls were recorded in analogue form onto reel-to-reel tapes (T. Struhsaker) and cassette tapes (R. Seyfarth and D. Cheney), and were later digitised at 22.05 kHz or 44.1 kHz with a 16 bit resolution. For a more

detailed description of recording equipment see (Struhsaker 1967b; Seyfarth et al. 1980b). South African vervet and green monkey calls were recorded by TP using a digital Marantz PMD661 solid-state recorder (44.1 kHz sampling rate, 16bits accuracy) connected to a Sennheiser ME66/K6 directional microphone.



**Figure 2.1** Distribution of African green monkeys (*Chlorocebus*) and sites at which recordings were made for analysis of call structure. Species distributions are shaded and modified from Lernould (1988) and Kingdon (1997).

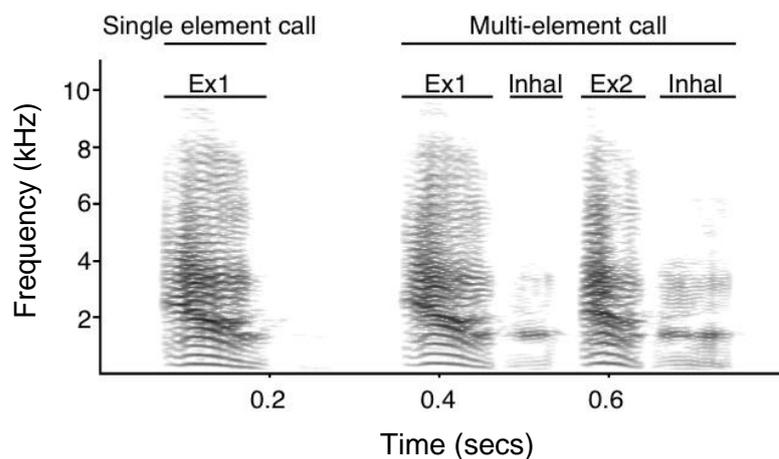
### Call selection

Bark vocalisations were frequently produced in long calling bouts. Consecutive bark elements (the basic units represented by a continuous sound) were identified using the pulse-train analysis of Avisoft SASLab Pro (version 5.1.17) and the start and end points of each element were recorded. From these labels, we calculated inter-unit intervals and used a log survivor function (Slater and Lester 1982) to determine a time threshold of 75 ms, below which elements were classified as belonging to the same call. On the basis of this, barks calls can be made up of one or more bark elements, with inter-call intervals exceeding all intra-call intervals and multi-unit barks frequently containing first exhalation (Ex1), within-call exhalation (Ex2) and inhalation (Inhal) call elements (definition of terms modified from Struhsaker 1967), see Figure 2.2.

It was not possible to analyse all calls from a calling bout, as males were often too far away from the microphone when they started calling and/or many calls were overlapped by the calls

of conspecifics. As such, call samples were taken from as near as possible to the start of each calling bout, but had to be tailored to the number and quality of calls recorded.

To investigate population-level differences in adult male bark vocalisations, we analysed barks from 12 green monkeys, 12 South African vervets, and 13 East African vervets. Unless insufficient calls of adequate quality were available, 10 bark calls were selected from the calling bout of each individual, resulting in a total call sample of 352 bark calls (120 green monkey barks, 120 South African vervet barks, and 112 East African vervet barks). To investigate individual-level differences, we analysed the bark vocalisations of six adult male South African vervets. We could not extract uninterrupted whole calls from all bouts of each individual, and as such we were not able to test temporal differences at the level of the whole call. To identify differences in element structure we selected 20 Ex1 bark elements from five calling bouts per male, resulting in a call sample of 100 bark elements per male and a total of 600 bark elements.



**Figure 2.2** Adult male bark series illustrating a single element and a multi-element call. Labels indicate temporal characteristics and different element types. The spectrogram was created using Avisoft SASLab Pro, with a 512 FFT and a Hamming window.

### Acoustic analysis

For analysis of population differences, call duration, Ex1 duration and the number of elements within a call were calculated based on the call labels described above (Figure 2.2). For analysis of population and individual differences, spectral analysis was carried out on Ex1 call elements only. Call processing prior to spectral analysis was carried out using Avisoft SAS Lab Pro. Calling bouts were first highpass filtered at 80 Hz to remove background noise below the lowest frequency of calls, following which, undisturbed Ex1 bark elements of high signal-to-noise ratio were extracted and padded with silent margins. Next, the frequency and temporal resolution of calls was adjusted to optimise measurement accuracy; for robust measures of energy distribution throughout the call unit, sampling frequency was reduced to 16 kHz, and calls were transformed using a fast Fourier transformation (FFT) size of 1024 points, Hamming

window and 93.75% overlap. These same settings were used to extract a measure of tonality, but because calls frequently only exhibited tonality at low frequencies and higher frequency noise hindered calculations, calls were first lowpass filtered at 1.2 kHz. For measures relating to fundamental frequency (F0), sampling frequency was reduced to 8 kHz and calls were transformed using an FFT size of 1024 points, Hamming window and 96.87% overlap. The resulting frequency-time spectra were analysed with LMA, a custom software sound analysis tool developed by KH, and parameters used for analysis are described in Table 2.1.

**Table 2.1** Description of the acoustic parameters used to describe the temporal and spectral structure of barks

| Measurement              | Description  |
|--------------------------|--|
| Temporal                 |  |
| Duration (ms)            | Duration of call   |
| Ex1 duration (ms)        | Duration of single or first call element                                   |
| Element number           | The number of elements per call  |
| Spectral                 |  |
| F0 (Hz)                  | Mean fundamental frequency across all time segments                        |
| F0 start (Hz)            | Fundamental frequency at the start of the call unit                        |
| F0 end (Hz)              | Fundamental frequency at the end of the call unit                          |
| F0 linear trend          | Factor of linear trend of fundamental frequency                            |
| Tonality (%)             | Percentage of tonal time segments for which F0 can be calculated           |
| First_quartile (Hz)      | Median first frequency quartile across all time segments                   |
| First quartile_1-4 (Hz)  | Mean first frequency quartile at 1st, 2nd, 3rd and 4th temporal quartiles  |
| Second_quartile (Hz)     | Median second frequency quartile across all time segments                  |
| Second quartile_1-4 (Hz) | Mean second frequency quartile at 1st, 2nd, 3rd and 4th temporal quartiles |
| Third_quartile (Hz)      | Median third frequency quartile across all time segments                   |
| Third quartile_1-4 (Hz)  | Mean third frequency quartile at 1st, 2nd, 3rd and 4th temporal quartiles  |
| Frequency range (Hz)     | Mean frequency range   |
| Peak frequency (Hz)      | Median peak frequency across all time segments                             |
| Peak frequency_1-4 (Hz)  | Mean peak frequency at 1st, 2nd, 3rd and 4th temporal quartiles            |
| PF linear trend          | Factor of linear trend of peak frequency                                   |
| PF deviation (Hz)        | Mean deviation between peak frequency and linear trend                     |

## Statistical Analysis

To assess population and individual differences in male bark calls, we first applied a stepwise method to identify a subset of optimum variables for each classification. We set population or caller identity as the grouping variable and entered all temporal and spectral parameters into a stepwise variable selection using the stepclass function of the R package

"klaR" (Weihs et al. 2005) with leave-one-out cross-validation. To assess the degree to which barks could be correctly assigned and to determine which structural properties contributed most to differentiating between the different populations or callers, we then entered the selected variables into a linear discriminant analysis (LDA) using the `lda` function of the R package "MASS" (Venables and Ripley 2002) with a jack-knife leave-one-out method. We compared the classification results of the LDA to those of a nested permuted discriminant function analyses (pDFA; Mundry and Sommer 2007) to control for the pseudoreplication introduced by using multiple calls from a single calling bout. All statistical analyses were carried out using R (R Development Core Team 2011), and before running tests we ensured that assumptions were met (see S2.1-S2.2 in supplementary).

## Results

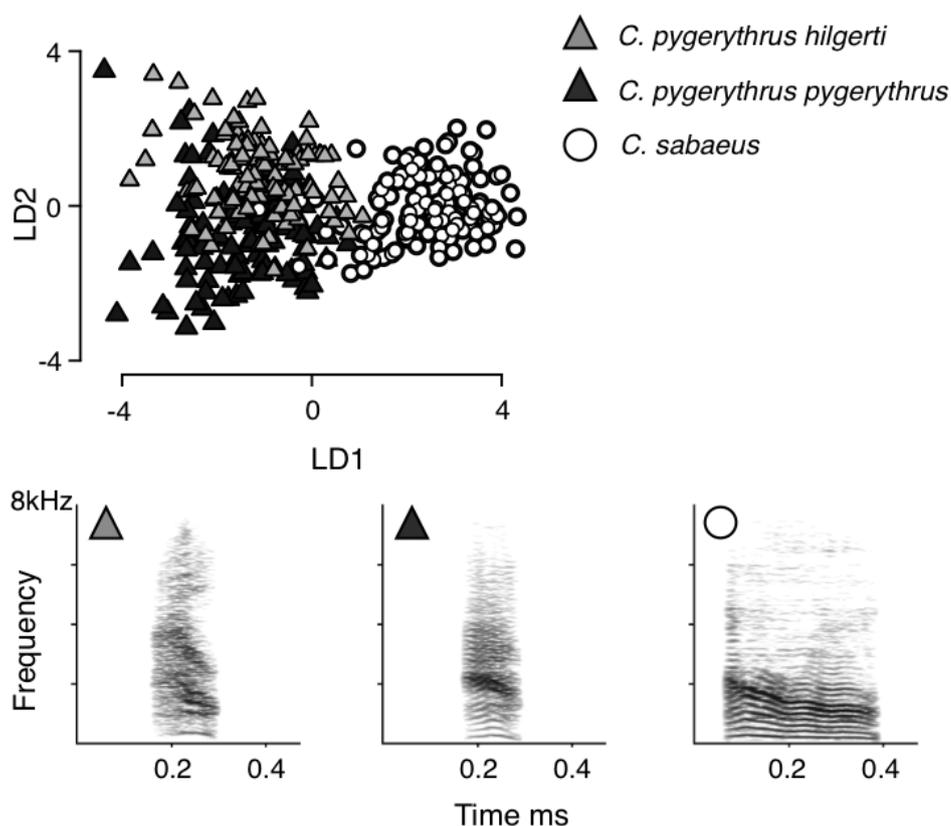
Stepwise variable selection identified a subset of 5 acoustic parameters that best differentiated between calls from different populations. These parameters were: Ex1 duration, F0start, F0 linear trend, frequency range, and PF deviation. On the basis of differences in these parameters, LDA (using a leave-one-out method) correctly classified 82% of bark units to their population of origin; this result was supported by the pDFA, which also correctly classified 82% of calls. Calls were most distinct at the species level, with 96% of calls being assigned to the correct species, compared to 77% of vervet calls that were assigned to the correct subspecies (Table 2.2).

**Table 2.2** Percentage of calls assigned to each population and descriptive statistics (mean  $\pm$ SD) of acoustic parameters used for classification

|              | Call assignment |              |             | Acoustic parameters |              |                 |                 |               |
|--------------|-----------------|--------------|-------------|---------------------|--------------|-----------------|-----------------|---------------|
|              | Green monkey    | South Vervet | East Vervet | Ex1 duration        | F0 start     | F0Linear trend  | Frequency range | PF deviation  |
| Green monkey | 94 %            | 3 %          | 3 %         | 262 $\pm$ 110       | 254 $\pm$ 40 | -0.14 $\pm$ 0.1 | 2197 $\pm$ 708  | 128 $\pm$ 68  |
| South Vervet | 1 %             | 78 %         | 21 %        | 99 $\pm$ 16         | 282 $\pm$ 57 | -0.11 $\pm$ 0.2 | 3116 $\pm$ 831  | 101 $\pm$ 53  |
| East Vervet  | 4 %             | 22 %         | 74 %        | 113 $\pm$ 19        | 320 $\pm$ 61 | -0.10 $\pm$ 0.2 | 3565 $\pm$ 645  | 156 $\pm$ 115 |

The first discriminant function separated green monkey from vervet barks and accounted for 90% of the total variance explained. Ex1 duration contributed most to this classification, with West African green monkeys typically producing longer barks than vervet monkeys, although the Ex1 duration of green monkey calls was also quite variable. The second discriminant function

accounted for 10% of the total variance explained, and separated South and East African vervet calls. This discriminant function was most dependent on differences in F0 start, Ex1 duration and frequency range (Table 2.2, Figure 2.3).



**Figure 2.3** Scatterplot and spectrograms illustrating population differences in the acoustic structure of *C. sabaesus*, *C. p. hilgerti*, and *C. p. pygerythrus* barks. The scatterplot presents the distribution of the first and second LDA discriminant scores. Spectrograms illustrate a typical call exemplar for each call group, with typical calls defined as those that were most likely to be assigned by LDA to the correct caller/population. Spectrograms were made with a 512 FFT and a Hamming window

Stepwise variable selection carried out to look at individual differences in the acoustic structure of Ex1 call elements identified 5 acoustic parameters to enter into a subsequent LDA. These parameters were F0, first quartile\_1, first quartile\_2, first quartile\_4 and Ex1 duration. Entering these parameters into LDA (using a leave-one-out method) resulted in 70% correct assignment (range=59-94%) of South African vervet barks, very similar to the 71% correctly assigned with pDFA. The first discriminant function described 74% of the total variance explained and was most influenced by F0. The second discriminant function described 17% of the total variance explained and was most influenced by measures of the first quartile energy band.

## **B. Behavioural responses to bark vocalisations**

### **Experimental protocol**

Playback experiments were conducted with South African vervets between January and June 2012 by TP and ON. Study subjects were 11 habituated and individually recognised adult male South African vervets from four free-ranging groups within Loskop Dam Nature Reserve in South Africa. For each experiment we played back bark calls elicited by a leopard model. These barks originated from own-group adult males ("South-own"), unknown adult males of the same subspecies ("South-unknown"), and adult male green monkeys ("West"). Barks used as stimuli for the West and South-own playback conditions were selected from recordings of green monkeys and South African vervets that were used in the structural analysis of bark calls. Barks used as stimuli for the South-unknown playback condition were recorded by Daniel van der Post from free-ranging groups in the Okavango Delta (18°25'42S; 21°53'39E) and Limpopo (22°54'25S; 28°2'28E) both in Botswana, and from Lajuma Research Centre (23°2'21S; 29°26'58E) in South Africa. Playback stimuli were made up of 6-12 bark units with a high signal-to-noise ratio that were produced as single and double unit exhalations sometimes interspersed with inhalation units. All bark units were taken from a single calling bout, with inter-call intervals, call compositions and maximum amplitude held constant between playback conditions. Reflecting temporal differences identified in the above section of this study, however, mean duration of call elements was longer in West playback stimuli than in South-unknown and South-own stimuli.

Barks were played back to male vervets using a within-subjects design such that, with one exception, each of the eleven subjects experienced one playback of each condition. This exception was due to a male migration which left one study subject as the sole male of his group before a final experiment could be carried out. We therefore conducted a total of 32 playback experiments (11 West, 11 South-unknown and 10 South-own). The order of playback trials was balanced across conditions, and to avoid habituation effects playback experiments were carried out on one group with a minimum separation of six days. To avoid pseudo-replication, a different playback sequence was used for each playback experiment. As much as possible each call sequence was produced by a different individual, and no individual contributed calls to more than 2 playback stimuli.

Playbacks were initiated when the study subject was sitting resting on the ground or low down (<2m) in a tree, and when the caller (in South-own condition) or another adult male (South-unknown and West conditions) was out of sight. This was to control for the subject responding to observed male presence rather than to identity cues present in the broadcast call.

Playbacks were not carried out within 60 minutes of the natural production of bark calls or any other alarm calling within earshot of the experimenter. In addition, since we were interested in the propensity of the subject to climb up into a tree, it was necessary that at the time of the experiment there was a tree of >2m in close vicinity to the subject. Prior to an experiment, the loudspeaker was hung using a net bag behind a natural obstacle at 1-2m from the ground and 42-59m from the study subject. Playback stimuli were broadcast using a Marantz PMD-661 solid-state recorder connected to a loudspeaker (David Active, VISONIK, Berlin), with maximum amplitude set within the range of natural calling behaviour (66-79dB at 10m from source, measured using a Voltcraft 322 sound level meter). Experiments were discarded if the subject moved prior to stimulus presentation (n=1) if there were technical problems with the equipment (n=2), if the individual was lost before the end of the experiment (n=1), or if the subject's behaviour was altered by the presence of human food (n=1).

### **Behavioural analysis**

Following each playback, subjects were filmed for at least 30seconds using a Sony Handycam (DCR-HC90E). At the end of these 30 seconds we recorded whether the subject had climbed more than 2m up into a tree, and, using a tape measure, the maximum horizontal distance travelled. To assess behavioural effects over a longer time period, subjects were followed for 30 minutes (from the point at which the playback was broadcast), and at each 5-minute interval we recorded their height (as being more or less than 2m from the ground), and their position using a handheld GPS (Garmin GPSMAP 60CSx). We additionally took GPS points of the position of the loudspeaker, the subject's initial position at time of playback, and of the subject's position 3 minutes after the playback experiment.

Post-experiment we used GPS points to calculate the distance that a subject moved relative to the loudspeaker within the first 3 minutes (distance from speaker at 3 minutes, minus distance from speaker at start), and the amount of time an individual spent at more than 2 m from the ground within 30 minutes of the playback (using height measures taken at each 5-minute interval for each experiment). Videos were imported into Adobe Premiere Pro CS4 with a time resolution of 25 frames/second, and frame-by-frame analysis of videos was used to score the duration of the subject's first orientation towards the speaker. Because video encoding is susceptible to observer-bias, 50% of the videos were reanalysed by a second condition-naive observer. Inter-observer reliability demonstrated moderate agreement (intra-class correlation coefficient 0.7). A description of behavioural measures is given in Table 2.3.

**Table 2.3** Description of the behavioural measures used to describe subjects' responses to playback experiments

| Behavioural measure            | Description   |
|--------------------------------|---|
| Strength of response:          |   |
| First orientation (s)          | Duration of first orientation towards loud speaker                |
| Initial displacement (m)       | Maximum distance travelled within 30 s of experiment              |
| Leopard-appropriate response   |   |
| Height_30s                     | Does subject climb to >2m within 30 seconds of experiment         |
| Height_30mins                  | Is subject >2m high within the 30minutes following experiment     |
| Male-Male competitive response |   |
| Loudspeaker_fast approach (m)  | Distance approached towards loudspeaker 3 minutes post-experiment |
| Loudspeaker_min approach (m)   | Minimum distance to loudspeaker within 30 minutes of experiment   |

### Statistical Analysis

To test whether bark origin would have an effect on the strength of response we used a general linear mixed model (GLMM) with Gaussian error structure to assess differences in the duration of subjects' first orientation towards the playback speaker, and a GLMM with Poisson error structure to assess differences in initial displacement. We ran GLMMs with binomial error structure to test the effect of bark origin on leopard-typical response behaviours, more specifically looking at differences in whether the male climbed immediately up into a tree (height\_30s) and whether the male spent more time up in a tree over the next 30 minutes (height\_30mins). Lastly, to test whether bark origin would have an effect on the male-male aggressive behaviours we used general linear mixed models (GLMM) with Gaussian error structure to analyse differences in the distance approached towards the speaker after 3 minutes (loudspeaker\_fast approach) and the minimum distance to loudspeaker over the next 30 minutes (loudspeaker\_min approach).

For all GLMMs, playback condition (West, South-own, South-unknown) and sequence order were entered as fixed effects and subject was entered as a random effect; to test for differences in subjects' height over the 30minute period, experiment was added as an additional random effect to account for interval data being included as separate data points. Maximum Likelihood was used to achieve more reliable P-values for models run using a Gaussian error structure. Likelihood ratio tests were used to compare full models with null models (comprising only of sequence order and the random effect), and Markov Chain Monte Carlo sampling using the functions `pvals.fnc` and `aovlmer.fnc` of the R package `languageR` (Baayen 2011) was applied to calculate p-values for the different levels for behaviours that differed between playback

conditions. All models were fitted in R using the function `lmer` of the R-package `lme4`. For details of test assumptions see S2.3 in supplementary.

## Results

The majority of playbacks (30/32) elicited an orienting response, with the male looking immediately in the direction of the speaker, whilst fewer playbacks (13/32) elicited immediate (within 30 seconds) displacement. With regards to the strength of response, there was a significant effect of playback condition on the males' first orientation towards the speaker. In response to barks of South-unknown origin, males' first orientation was significantly longer than first orientation towards barks of own group males (likelihood ratio test:  $-4.7 \pm 1.5$ ,  $t = -3.2$ ,  $P_{MCMC} < 0.01$ ), and first orientation towards green monkey barks (likelihood ratio test:  $-4.7 \pm 1.4$ ,  $t = -3.3$ ,  $P_{MCMC} < 0.01$ ). There was no significant effect of playback condition on males' initial displacement, or on leopard-typical response behaviours. Looking at male-male aggressive behaviours, playback condition did not affect whether subjects immediately approached the loudspeaker (loudspeaker\_fast approach), but there was a significant effect on loudspeaker\_min approach, with males approaching closer in response to South-unknown stimuli ( $29 \pm 15$ m) than to South-own ( $53 \pm 18$ m, likelihood ratio test:  $21.4 \pm 5.8$ ,  $t = 3.7$ ,  $P_{MCMC} < 0.01$ ), or West African calls ( $42 \pm 22$ m, likelihood ratio test:  $13.4 \pm 5.6$ ,  $t = 2.4$ ,  $P_{MCMC} < 0.05$ ). For all behavioural measures, descriptive statistics and results from GLMMs are presented in Table 2.4.

**Table 2.4** Description of response behaviours (RB: mean $\pm$ SD), occurrence of behaviour (Occ: measured out of total experiments or for Height\_30m out of all incidences of data collection which was six/experiment at 5 minute intervals), and results from general linear mixed models for the six behavioural variables measured during playback experiments.

| Behaviour                        | Own            |       | Unknown         |       | Green          |       | $\chi^2$ | df | P               |
|----------------------------------|----------------|-------|-----------------|-------|----------------|-------|----------|----|-----------------|
|                                  | RB             | Occ   | RB              | Occ   | RB             | Occ   |          |    |                 |
| First orientation (s)            | 2.1 $\pm$ 2.0  | 8/10  | 5.3 $\pm$ 2.9   | 11/11 | 1.8 $\pm$ 1.6  | 11/11 | 11.92    | 2  | <b>&lt;0.01</b> |
| Initial displacement (m)         | 1.0 $\pm$ 1.6  | 5/10  | 3.8 $\pm$ 7.1   | 4/11  | 1.9 $\pm$ 4.2  | 4/11  | 2.19     | 2  | 0.33            |
| Height_30s                       | NA             | 3/10  | NA              | 3/11  | NA             | 2/11  | 0.38     | 2  | 0.83            |
| Height_30mins                    | NA             | 35/60 | NA              | 43/66 | NA             | 28/66 | 3.03     | 2  | 0.22            |
| Loudspeaker_FA (m) <sup>+</sup>  | 5.1 $\pm$ 10.4 | 1/10  | -2.5 $\pm$ 15.0 | 4/11  | 0.5 $\pm$ 11.9 | 3/11  | 2.14     | 2  | 0.34            |
| Loudspeaker_MA (m) <sup>++</sup> | 53 $\pm$ 18    | NA    | 29 $\pm$ 15     | NA    | 42 $\pm$ 22    | NA    | 11.15    | 2  | <b>&lt;0.01</b> |

<sup>+</sup>FA = fast approach, negative values represent an approach towards the speaker

<sup>++</sup>MA = min approach

## Discussion

### Population differences in call structure

Despite their overall acoustic similarity, the bark calls of West African green monkeys, East African and South African vervets could be distinguished on the basis of spectral and temporal differences in call structure. Of the three populations, correct classification at the species level was high (96% of calls) indicating that call differences were larger between than within species. The calls of vervet subspecies were also distinguishable, although to a lesser degree, being correctly classified in 77% of cases, resulting in an overall correct classification of 82%. This is in accordance with genetic data which identifies green monkeys as especially distinct from other *Chlorocebus* taxa (Haus et al. 2013) with an estimated divergence time of 2.76-2.81MYA, compared to 1.59-1.48MYA divergence of vervet subspecies (Wertheim and Worobey 2007). Thus this study indicates that in some primate taxa, calls produced in alarm contexts can also be used to distinguish between species. A similar pattern is found at the within-species level in the short call of the pika (*Ochotona princeps*), which exhibits variation in duration and F0 between geographically isolated but not more neighbouring populations (Conner 1982). In contrast the alarm calls of western gray and eastern rufous mouse lemurs (*Microcebus murinus* and *M. rufus*), despite diverging 6.7-7.3MYA (Weisrock et al. 2012), do not differ (Zimmermann et al. 2000). One explanation for these differences could be that African green monkey barks and pika short calls are both also produced outside of alarm contexts (Struhsaker 1967b; Svendsen 1979), they may therefore be less constrained than the more stimulus-specific mouse lemur alarms.

With regards to how African green monkey barks differ, this study identified the duration of Ex1 bark elements as the most influential parameter distinguishing between green monkey and vervet barks, with green monkeys producing longer, but also lower frequency, call elements than vervets. East African vervets on the other hand, produced barks with a lower fundamental frequency and a larger frequency range than their South African sister taxa. The production of calls with lower frequency bands and lower F0 could well be the result of a larger body size (and corresponding larger vocal anatomy), a correlation found to hold across a wide range of primate species (Hauser 1993a). Support for this explanation comes from a comparative study of cranial measurements across the genus which identifies green monkey samples as being larger than those of vervets (Elton et al. 2010). Size differences between South and East African vervets are, however, unclear (Pasternak et al. 2013). The prolonged bark produced by male green monkeys could also be the result of a larger body size in relation to lung capacity (Fitch and Hauser 2002), but it seems unlikely that this could account for such large differences in call duration. One other

suggestion is that male barks function as a sexual display as well as an alarm call (Price et al. in review), in which case the longer bark duration could be a sexually selected trait.

Importantly, while African green monkeys exhibit geographic variation in the subtle acoustic structure of their barks, the long calls of leaf monkeys (*Presbytis*) and the song of crested gibbons (*Nomascus*) display geographic differences in call type and sequential order (Thinh et al. 2011; Meyer et al. 2012). This higher level of flexibility cannot be explained solely by longer divergence times (Thinh et al. 2010; Meyer et al. 2011); instead it may be the result of differences in call function, in accordance with the prediction that alarm calls should diverge more slowly than calls under sexual selection as a result of stronger stabilising and weaker directional selective pressures (Struhsaker 1970). Thus the fact that barks are produced in response to predators and during aggressive interactions could account for it being more flexible than mouse lemur alarms, but more constrained than leaf-monkey and crested gibbon mating calls.

### **Behavioural responses to bark vocalisations**

The duration of the subject's first orientation towards the speaker was significantly longer in response to unknown conspecific calls than to known conspecific calls, strongly suggesting that adult male vervets differentiate between the calls of known and unknown males and are more attentive to the presence of "strangers". That males differentiate between these call categories is supported by the earlier finding of this study that male barks exhibit individual differences in call structure. This result does not, however, require that males recognise group members individually (i.e. "true" individual recognition; Tibbetts and Dale 2007). That males are more attentive to unknown calls is not surprising, as vervet monkeys live in relatively stable multi-male multi-female groups (Fedigan and Fedigan 1988) and actively defend their territory against intruding males (Struhsaker 1967b). Additionally, subjects' longer orientation to unknown conspecific barks than to unknown inter-specific barks suggests that inter-specific differences in this call are a socially relevant cue. We predicted that subjects should respond to all barks with the leopard-typical response of climbing up into a tree. We found that subjects climbed into a tree on hearing the playback stimulus in 8/32 experiments, with no effect of playback condition. Thus our results show that conspecific and inter-specific barks elicit leopard-typical escape responses, demonstrating some degree of the predicted constraint that receivers should be under to respond appropriately to alarm calls. Nevertheless, the finding that arboreal escape responses were elicited in only 25% of trials suggests that bark calls may be a less urgent alarm call than the female alarm chirp (which in green monkeys caused subjects to climb into

trees in 50% of the experiments, Price and Fischer 2013), and that at least for adult males, context may play an influential role in the process of meaning attribution. Interestingly, on hearing the playback stimulus males also approached the playback speaker in 25% of trials, suggesting that barks may also function as an intra-sexual display. In addition, males were more likely to move nearer towards the playback speaker in the 30 minutes following experiments after hearing the calls of an unknown conspecific than after hearing the calls of a known conspecific or interspecific male, corroborating the assumption that males are able to differentiate between the calls of own-group and unknown males and are more attentive to the calls of unknown males of the same species.

## **Conclusion**

Subtle variation in the structure of the African green monkey bark can be used to differentiate between species and to a lesser extent subspecies. This strengthens the general consensus that genetic differences account for a large degree of the vocal flexibility found within and between primate species, not only in loud calls but also in calls produced in alarm contexts. Nevertheless, the study does support Struhsaker's (1970) early claims that, within the *Chlorocebus* genus, alarm vocalisations exhibit fine-scaled temporal differences but are otherwise relatively stable, and that calls produced in alarm contexts are under stronger stabilising selection than calls functioning primarily in mate choice and/or male-male competition. From the perspective of bark perception, this study suggests that adult males differentiate between the barks of own-group and unknown males, and respond more strongly to unknown conspecific than unknown heterospecific barks. Thus within and between-species variation in bark structure appears to be socially relevant. That males respond to barks both by climbing into the trees and by approaching the speaker suggests that adult male barks may have a dual function; a terrestrial alarm call and a male display. More generally this study supports the hypothesis that, in contrast to human speech, flexibility in primate vocalisations is shaped by call function.

## **Acknowledgements**

We thank Mpumalanga Parks Board for permission to work at Loskop Dam Nature Reserve, and the *Diréction des Parcs Nationaux* and *Ministère de l'Environnement et de la Protéction de la Nature de la République du Sénégal* for permission to work in Niokolo-Koba National Park. We are also grateful to Robert Seyfarth, Dorothy Cheney, Tom Struhsaker and Daniel van der Post for kindly providing vervet calls for inclusion within this study; to Jannie and Elizma Coetzee, Tersia de Beer and Alan Barrett for their logistical support while working in South Africa; and to Jacky Bassene, Cheikh Sane, Samba Ciss, Nele Fröse and Matthias Klapproth for their assistance with data collection in Senegal. Thanks also go to Mathieu Cantat for his assistance with video analysis, to Christof Neumann for statistical advice, and to Tanja Haus and Gisela Kopp for their comments on the manuscript. The study was supported by funding from the *Deutscher Akademischer Austauschdienst*, The Leakey Foundation and The Wenner-Gren Foundation.

## **Ethical standards**

The experiments described in this paper comply with the current laws of Senegal, South Africa and Germany, and were conducted with the agreement of the Mpumalanga Parks Board, the *Diréction des Parcs Nationaux*, and the *Ministère de l'Environnement et de la Protéction de la Nature de la République du Sénégal* (Permit numbers: 0383/24/03/2009; 0373/10/3/2012).

## **Conflict of interest**

The authors declare that they have no conflict of interest.

## Supplementary

### S2.1 Geographic differences in call structure

Before running LDA to assess how call structure varied between populations we visually inspected histograms of the 31 acoustic parameters to check that there were no outliers and that their univariate distributions were approximately symmetrical. Predictor variables that were not symmetrically distributed were transformed. When taking the log or square-root of the parameter did not achieve symmetrical distribution, we used the Box-Cox function (R package "mass", Venables and Ripley 2002) to find the optimum transformation based on maximum-likelihood methods. After transformation all acoustic parameters were standardised using a Z-transformation in order to make them more comparable in terms of their units of measure and variance. See Table S2.1 for which transformation was applied to each parameter.

**Table S2.1** Transformations applied to the raw data of acoustic parameters prior to z-transformation and LDA

| Parameter         | Transformation |
|-------------------|----------------|
| Duration          | Log            |
| Element number    | None           |
| Ex1 duration      | $\wedge -0.8$  |
| F0                | None           |
| F0start           | None           |
| F0end             | $\wedge -0.1$  |
| F0 linear trend   | None           |
| Tonality          | +3 then log    |
| First quartile    | None           |
| First quartile_1  | None           |
| First quartile_2  | None           |
| First quartile_3  | None           |
| First quartile_4  | None           |
| Second quartile   | None           |
| Second quartile_1 | None           |
| Second quartile_2 | None           |
| Second quartile_3 | None           |
| Second quartile_4 | Log            |
| Third quartile    | None           |
| Third quartile_1  | None           |
| Third quartile_2  | None           |
| Third quartile_3  | None           |
| Third quartile_4  | None           |
| Frequency range   | None           |
| Peak frequency    | None           |
| Peak frequency_1  | None           |
| Peak frequency_2  | None           |
| Peak frequency_3  | None           |
| Peak frequency_4  | None           |
| PFlinear trend    | None           |
| PFdeviation (Hz)  | Log            |

To check the multivariate assumptions of LDA we ran a MANOVA, with population as the independent variable and the Z-scores of spectral parameters as dependant variables. We visually inspected plots of residuals to assess homogeneity and normality, and plotted mahalanobis distance to assess multivariate normality; data indicated no obvious deviations from these assumptions.

## S2.2 Individual differences in call structure

Before running LDA we carried out the same tests as described above to ensure that the 29 parameters met univariate and multivariate assumptions. Table S2.2 gives details on transformations applied prior to analyses.

**Table S2.2** Transformations applied to the raw data of acoustic parameters prior to z-transformation and LDA

| Parameter         | Transformation      |
|-------------------|---------------------|
| E x1 duration     | Log                 |
| F0                | None                |
| F0start           | None                |
| F0end             | None                |
| F0 linear trend   | +3 then square root |
| Tonality          | +3 then log         |
| First quartile    | None                |
| First quartile_1  | $^2$                |
| First quartile_2  | None                |
| First quartile_3  | None                |
| First quartile_4  | None                |
| Second quartile   | $^{-1}$             |
| Second quartile_1 | None                |
| Second quartile_2 | $^{-0.8}$           |
| Second quartile_3 | None                |
| Second quartile_4 | None                |
| Third quartile    | None                |
| Third quartile_1  | None                |
| Third quartile_2  | None                |
| Third quartile_3  | None                |
| Third quartile_4  | None                |
| Frequency range   | None                |
| Peak frequency    | $^{1.9}$            |
| Peak frequency_1  | $^{-0.6}$           |
| Peak frequency_2  | None                |
| Peak frequency_3  | None                |
| Peak frequency_4  | None                |
| PFlinear trend    | +8 then $^{1.8}$    |
| PFdeviation (Hz)  | +3 then log         |

### S2.3 Behavioural responses to bark vocalisations

Before running General Linear Mixed models we visually inspected boxplots and histograms of non-binomial behavioural responses to assess heterogeneity of variance between conditions, approximate symmetric distribution and absence of outliers. Following this we fitted models (minus random effects) using the function `glm` (applying the transformations and GLMM family described in table S3), and determined variance inflation factors (VIFs) using the `vif` function of the R package `car` (Fox and Weisberg 2011), and visually assessed graphs illustrating the distribution of model residuals. VIFs indicated that collinearity was not an issue (largest VIF = 1.05). Variables run with gaussian error structure met the assumptions of normally distributed and homogenous residuals. Maximum distance travelled by the subject within 30 seconds did not meet these assumptions and we ran this model using a poisson error structure. Original data were overdispersed so we transformed these measures to ordinal data. Following this, data was still somewhat overdispersed (dispersion parameter = 2.1), but as overdispersion can lead to tests becoming anti-conservative, results indicating a non-significant difference (as we found in this case) are still reliable.

**Table S2.3** Transformations applied to the raw data of behavioural measures prior to statistical analysis, and family of GLMM applied

| GLMM   | Transformation | GLMM family |
|--|----------------|-------------|
| Duration of first look towards speaker           | Square root    | Gaussian    |
| Maximum distance within 30 seconds               | Ordinal        | Poisson     |
| Subject's height within 30 seconds               | None           | Binomial    |
| Subject's height over 30 minutes                 | None           | Binomial    |
| Proximity to loud speaker after 3 minutes        | None           | Gaussian    |
| Minimum distance to loud speaker over 30 minutes | None           | Gaussian    |



## Chapter 3

### **Vervets revisited: A quantitative analysis of alarm call structure and context specificity**

Tabitha Price<sup>1,2,3\*</sup>, Philip Wadewitz<sup>1,4</sup>, Thomas T. Struhsaker<sup>5</sup>, Dorothy L. Cheney<sup>6</sup>,  
Robert M. Seyfarth<sup>7</sup>, Julia Fischer<sup>1,2</sup>

<sup>1</sup> Cognitive Ethology Lab, German Primate Center, Germany

<sup>2</sup> Courant Research Centre for the Evolution of Social Behaviour, Georg August University of  
Göttingen, Germany

<sup>3</sup> Applied Behavioural Ecology and Ecosystem Research Unit, UNISA, South Africa

<sup>4</sup> Theoretical Neurophysics, MPI for Dynamics and Self-Organisation, Germany

<sup>5</sup> Evolutionary Anthropology, Duke University, United States

<sup>6</sup> Department of Biology, University of Pennsylvania, United States

<sup>7</sup> Department of Psychology, University of Pennsylvania, United States

In preparation for submission

### Abstract

**Background:** Ever since the alarm calls of the vervet monkey (*Chlorocebus pygerythrus*) were described as being predator-specific and found to elicit appropriate response behaviours in conspecifics, functionally referential communication has been centre stage in the quest to understand how language evolved. More than thirty years on, we revisit the vervet alarm calling system and use comprehensive quantitative analysis to re-assess the degree to which these calls are structurally discrete and context specific, thereby fulfilling one of the two classic criteria for functional reference.

**Results:** Cluster analyses show that at the level of the call elements, vervet alarm calls, especially those of the male, demonstrate a graded call structure and are not obviously separable into discrete call types. Nevertheless, within this graded system, discriminant function analysis of spectral characteristics does distinguish among alarm call elements with regard to the predator class, although more so within female than male calls. Female and male alarm call elements overlap in acoustic structure with call elements produced in non-alarm contexts indicating that context specificity is reduced when calls from other contexts are considered. General linear mixed models reveal that, at the level of the entire call, temporal characteristics differ with respect to predator type, but that these differences are also less obvious when alarm calls are compared to calls produced outside of the predator context.

**Conclusion:** At the level of call elements, vervet alarm calls constitute a graded system, especially in males, with limited specificity to particular predators when compared with calls from non-predator contexts. This may have created a selective pressure to produce the combinations of call elements that elicit predator-appropriate responses from conspecifics. Our findings support the view that call structure in non-human primates has a strong evolutionary component, with higher flexibility at the level of call usage. Thus, similarity with true referential communication is restricted to the cognitive capacities of listeners that are able to interpret these calls and to factor in contextual cues when choosing different response strategies. Our findings support the view that functionally referential signalling systems provide little insight into how semanticity in language evolved at the production level.

## Background

Language is a uniquely human trait, but a common argument is that the evolving language faculty would have been more likely to co-opt pre-existing and pre-linguistic neural and behavioural mechanisms than to evolve entirely novel language-specific cognitive modules (Hauser et al. 2002; Zuberbühler 2003; Hurford 2003; Schoenemann 2005; Seyfarth et al. 2005; Fedurek and Slocombe 2011). This idea grew in prominence in 1980, following the description by Seyfarth et al. (Seyfarth et al. 1980a; Seyfarth et al. 1980b) of the alarm calling system of the vervet monkey, *Chlorocebus pygerythrus* (previously *Cercopithecus aethiops*), within which predator-specific calls were attributed semantic properties (Seyfarth and Cheney 1993). The typical textbook version of this study describes three discrete call types that are emitted in response to leopards, raptors and snakes, and which function as "words" for these predators in the sense that listening vervets hearing playbacks of the calls perform appropriate predator avoidance behaviours even in the absence of the predator itself (Snowdon 1990; Zuberbühler 2003; Cäsar and Zuberbühler 2012; Townsend and Manser 2012). In showing that animal calls can denote objects and events in the external environment, this study was taken to indicate that animal vocal repertoires can contain both affective and symbolic calls, as suggested earlier by Peter Marler (Marler 1977). This challenged the "motivational/emotional hypothesis" (the prevailing view since Darwin 1872), which states that call production in non-human animals (hereafter animals) reflects the signaller's internal state with continuous gradation between call structures; and introduced as an alternative the "referential hypothesis" which states that some calls are evoked by specific external events irrespective of the signaller's internal state with discrete and highly stimulus specific variation in call structure (Evans and Marler 1995). Thus according to the referential hypothesis, the building blocks for semantic reference were already present in the semantic calls of our primate ancestors.

In the 1990's it became more apparent that the cognitive mechanisms underlying call production in animals were still poorly understood (Cheney and Seyfarth 1992; Marler et al. 1992). Macedonia and Evans (Macedonia and Evans 1993) therefore coined the term "functionally referential" to reflect the perspective that animal signals are "not exactly like human words, but rather appear to function in the same way" (Hauser 1997 p. 509). The referential hypothesis of call production was retained, however, with animal vocalisations described as falling along a continuum; from calls primarily reflecting the signaller's motivational state to calls reliably elicited by an external stimulus and unaffected by the signaller's motivational state (Macedonia and Evans 1993). Within this theoretical framework, Macedonia and Evans described criteria for classifying animal signals as functionally referential rather than

affective. The production criterion is that "referential signals should exhibit a degree of stimulus specificity" (Macedonia and Evans 1993 p179), (i.e. all or nearly all eliciting stimuli belong to a common category); the perception criterion is that the signal should bring about the same response as the eliciting stimuli even in the absence of any contextual cues (Marler et al. 1992; Macedonia and Evans 1993). The alarm, food associated, and social calls of many species of primate, other mammal, and bird have since been classified as functionally referential signals (reviewed in Zuberbühler 2003; Townsend and Manser 2012), and functional reference continues to be singled out as offering important insight into the evolution of symbolic communication in language (Zuberbühler 2003; Fedurek and Slocombe 2011; Townsend and Manser 2012).

It has, however, long been proposed that affect-based calls may also function referentially (i.e. meet the requirements of a functionally referential signal) if the signal receiver is able to interpret these signals to make inferences about external events (Premack 1975; Marler et al. 1992). Viewed in this way, the dichotomy between affect-based and referential calls disappears; the mechanisms underlying call production and perception are logically distinct; call production may or may not be related to the caller's internal state, and any similarity with true referential communication lies in how calls are perceived (Seyfarth and Cheney 2003b). To complicate matters further, it has been shown that animals may be able to categorise graded acoustic variation into different meaningful categories (Fischer 1998), indicating that distinctive calls are not a prerequisite for functionally referential communication (Fischer and Hammerschmidt 2001), and that the same vocalisations may vary both with regard to context and the caller's apparent emotional state (Fischer et al. 1995; Manser 2001).

More than forty years have passed since Struhsaker (Struhsaker 1967b) first described the vervet alarm calls, and more than thirty years have passed since the seminal study of Seyfarth et al. (Seyfarth et al. 1980a); yet it remains the best-known and most widely cited example of functional reference. This account is based on a combination of qualitative categorisation of alarm calls (Seyfarth et al. 1980b) and on quantitative analyses that identified structural differences between two alarm call types produced by adult females (Owren and Bernacki 1988), and between these female calls and acoustically similar non-alarm calls (Seyfarth et al. 1980b). However, these differences often consisted of variation in a single parameter at the level of a single individual. As such, we still lack a systematic and quantitative analysis of the structure of these alarm calls and similar calls in the repertoire, and we know relatively little about the alarm calling system of adult male vervets.

This study provides a quantitative assessment of the structure of vervet monkey alarm calls to evaluate the context specificity of these vocalisations. We analyse spectral structure at the

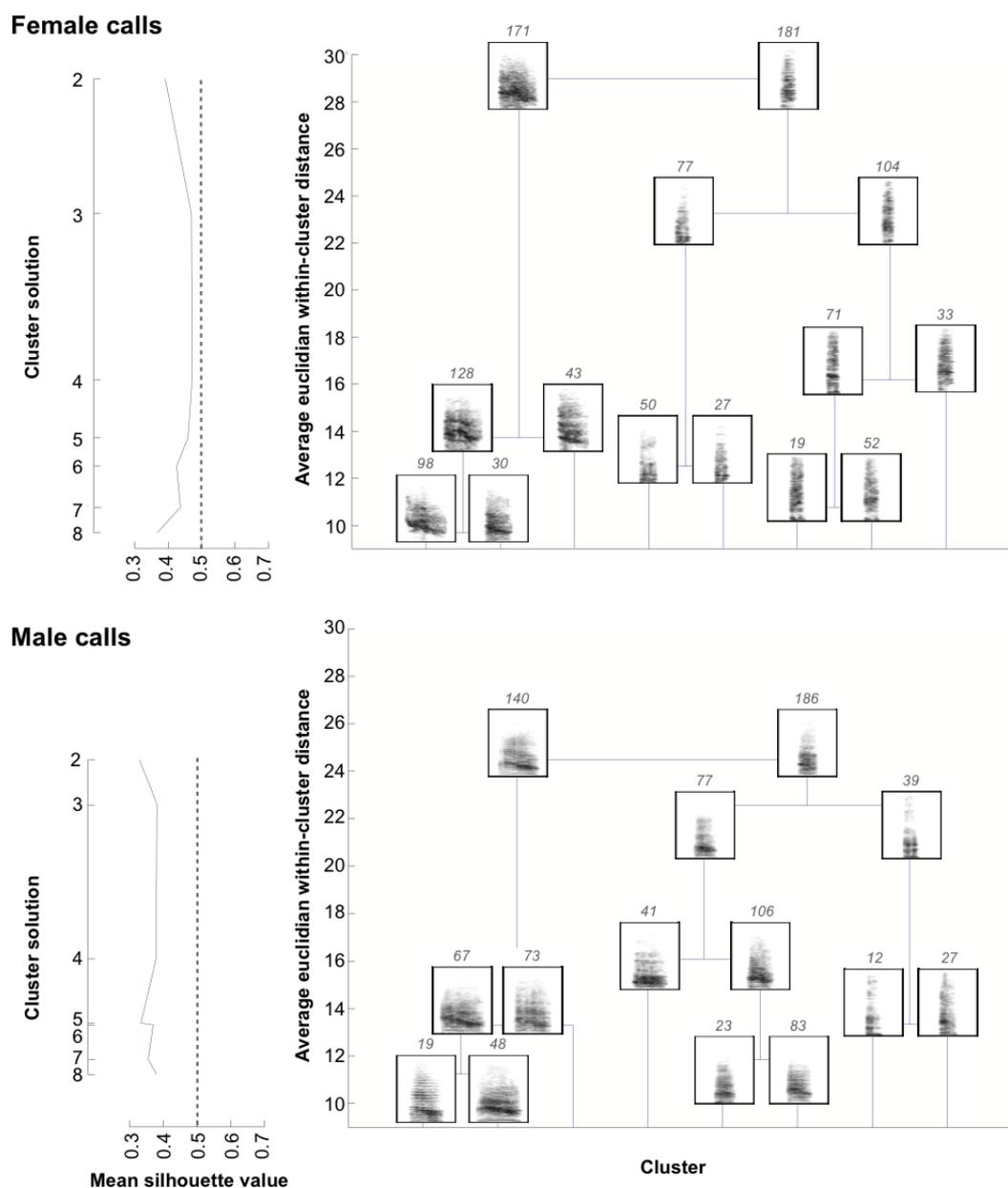
level of the call element (the basic unit of a vervet call, represented as a continuous tracing along the temporal axis of the spectrogram) and temporal structure at the level of the call (a call element produced singly, or a group of call elements). To investigate the degree to which the spectral structure of alarm call elements is discrete, we used cluster analysis to compare single acoustic elements drawn from the whole calls produced by vervets in response to leopards, raptors and snakes. To assess context specificity within the predator context, we applied discriminant function analysis to see how well these call elements could be distinguished on the basis of their spectral structure with respect to predator type. Following this, we ran general linear mixed models to determine whether predator type had an effect on temporal patterns of element production within whole calls. We then went on to apply discriminant function analysis and general linear models in the same way to a larger sample of isolated, single elements and whole calls produced in contexts of predator presence, and during social and aggressive interactions (including some calls recorded from a different population) in order to assess context specificity outside of the predator context. Because the calls of males and females differed substantially, we ran separate analyses for the two sexes, reducing the overall variation and allowing for a clearer picture within each sex.

## **Results**

### **A. Alarm call element structure**

Within k-means cluster analysis, silhouette coefficient values range from -1 to 1 and provide a quantitative measure of the stability of cluster solutions, with higher values indicating a better separation of clusters and therefore a more stable solution (Rousseeuw 1987). Silhouette coefficients above 0.5 are considered to indicate reasonable partitioning of data (Kaufman and Rousseeuw 2005). On the basis of this, k-means silhouette values calculated for vervet alarm call elements were low across all cluster solutions, with no clearly optimal cluster solution. Male call elements consistently exhibited lower silhouette values than female call elements, with a maximum silhouette co-efficient of 0.38 compared to a maximum silhouette co-efficient of 0.47 for female call elements (Figure 3.1). Thus results indicate that, in contrast to the traditional "three-call" categorisation originally made by ear and spectrographic analysis of naturally occurring calls, the single elements that make up the vervet's alarm call system, when considered separately, do not clearly subdivide into three categories. More generally, the low silhouette values suggest that, when taken as isolated individual units, vervet alarm calls (especially those of males) are not clearly separable into discrete call types at all but instead fall

below the 0.5 threshold and therefore constitute a graded call system. The acoustic structure of female and male alarm call elements is illustrated using hierarchical clustering of the same data set with spectrograms of elements drawn from whole calls (Figure 3.1).



**Figure 3.1** Classification of female (above) and male (below) alarm call elements using k-means and hierarchical cluster analysis. Line graphs (left) illustrate the mean silhouette values for 2-8 cluster solutions calculated using k-means cluster analysis. Silhouette values can vary from -1 to 1, with higher values indicating a better separation of the data. The dotted line at 0.5 represents the commonly accepted value for a reasonable partitioning of data. Dendrograms (right) illustrate the grouping of call elements using hierarchical cluster analysis; spectrograms depict a typical (showing smallest Euclidean distance to cluster centre) exemplar of each call group; all spectrograms were made with a 512 FFT and a Hamming window, and all spectrogram windows show a 0.2 second duration and a frequency range of 0 to 11025 Hz. The numbers written above each spectrogram indicate the number of call elements that belong to that call group.

## B. Context specificity

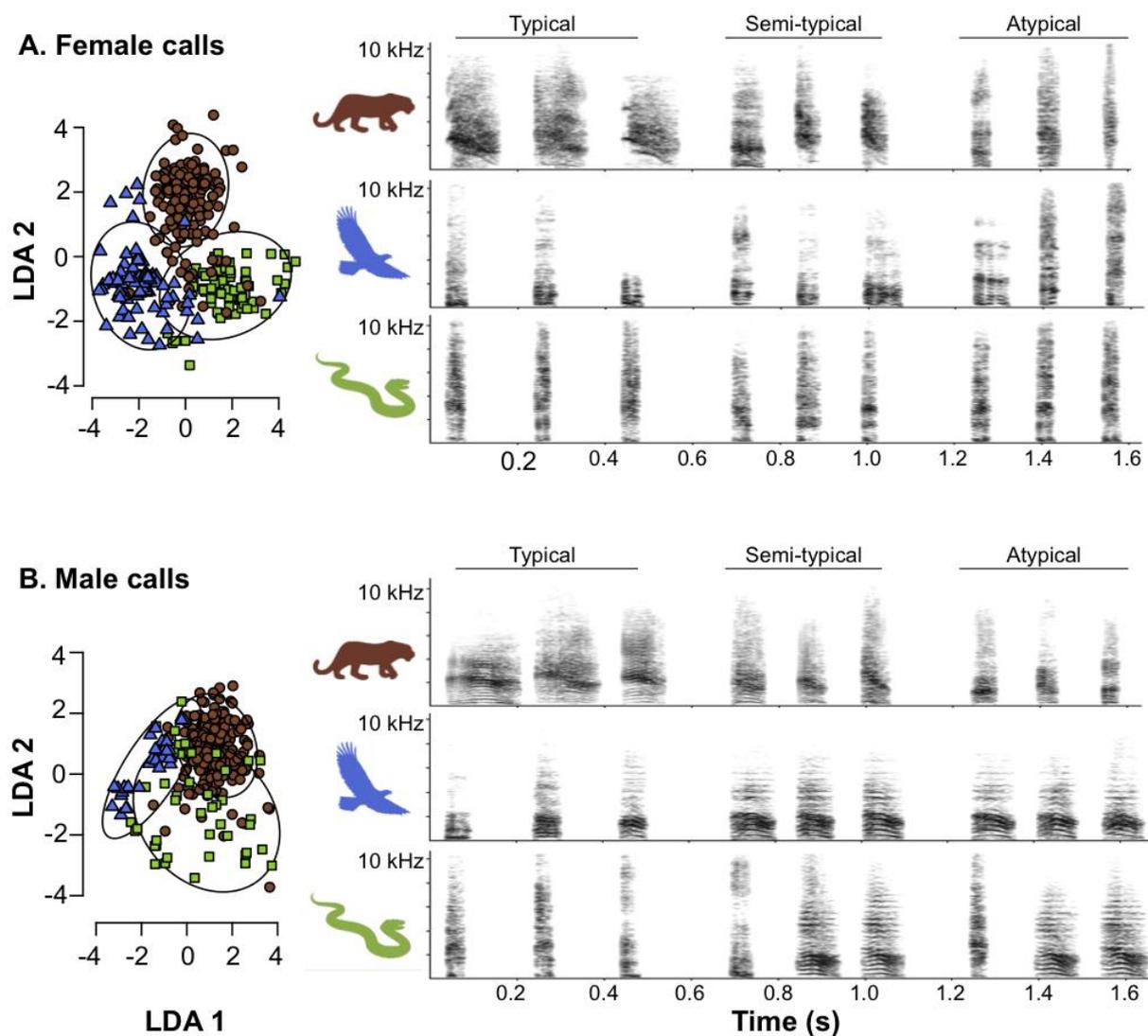
### Predator contexts

On the basis of differences in element duration, peak frequency, frequency range, PF jump, and PF linear trend, linear discriminant analysis (LDA) classified 90% of female call elements to the correct predator class. More specifically correct classification was high for all three predator classes with 87% of calls to leopards, 88% of calls to raptors and 95% of calls to snakes correctly classified (Figure 3.2a). A permuted discriminant function analysis (pDFA) carried out using the same variables on a subset of the call elements to control for individual identity correctly classified 89% of calls, thus supporting LDA findings. Element duration, peak frequency and frequency range were the most influential variables when discriminating between call elements given to different predators. The clusters identified by the three cluster solution of the k-means analysis largely corresponded to the three categories established by the LDA, and matched descriptions of call types from earlier studies by Struhsaker (1967b) and Seyfarth *et al* (1980b). In this study, long duration "chirp" elements were typically produced in response to leopards, low frequency "rraup" elements were typically produced in response to raptors, and broadband "chutter" elements were typically produced in response to snakes (Table 3.1). This matches Struhsaker's (1967b), and Seyfarth and colleagues' (1980b) earlier descriptions of "chirps", "rraups" and "chutters", with these calls produced predominantly (although not exclusively) in response to mammalian carnivores (including leopards, lions, and cheetah), raptors (usually martial eagles) and snakes (usually pythons), respectively.

**Table 3.1** Description (mean±SD) of acoustic parameters for female and male call elements at the level of the three cluster solution, and for female and male call elements produced in response to leopards, raptors and snakes.

| K-means three-cluster solution |                       |                 |                  |               |                 | LDA three class predator specificity |                       |                 |                  |               |                 |
|--------------------------------|-----------------------|-----------------|------------------|---------------|-----------------|--------------------------------------|-----------------------|-----------------|------------------|---------------|-----------------|
| Call group                     | Element duration (ms) | Peak freq (kHz) | Freq range (kHz) | PF jump (kHz) | PF linear trend | Call group                           | Element duration (ms) | Peak freq (kHz) | Freq range (kHz) | PF jump (kHz) | PF linear trend |
| Female                         |                       |                 |                  |               |                 | Female                               |                       |                 |                  |               |                 |
| Chirp                          | 105±30                | 2.3±0.5         | 4.0±1.1          | 1.0±0.7       | -1.5±1.1        | Leo <sup>+++</sup>                   | 97±36                 | 2.4±0.6         | 4.2±1.2          | 1.0±0.8       | -1.3±1.3        |
| Rraup                          | 42±16                 | 0.9±0.6         | 3.7±1.6          | 0.7±0.6       | -0.5±1.5        | Raptor                               | 42±17                 | 0.9±0.7         | 3.8±1.8          | 0.8±0.9       | -0.2±2.7        |
| Chutter                        | 36±14                 | 3.0±1.1         | 7.1±1.7          | 1.6±1.4       | 0.1±3.3         | Snake                                | 36±7                  | 3.0±1.1         | 7.4±1.2          | 1.8±1.4       | 0.1±3.5         |
| Male                           |                       |                 |                  |               |                 | Male                                 |                       |                 |                  |               |                 |
| Bark                           | 111±25                | 1.9±0.3         | 3.8±0.7          | 0.7±0.6       | -1.2±0.6        | Leo <sup>+++</sup>                   | 92±34                 | 1.9±0.4         | 3.8±0.8          | 0.6±0.4       | -1.1±0.8        |
| R-B <sup>+</sup>               | 65±21                 | 1.6±0.6         | 3.1±0.9          | 0.6±0.4       | -0.5±0.9        | Raptor                               | 68±24                 | 0.9±0.6         | 2.1±0.4          | 0.7±0.4       | 0.0±0.8         |
| C-B <sup>++</sup>              | 45±17                 | 1.3±1.0         | 4.8±1.5          | 1.5±1.0       | -2.4±3.4        | Snake                                | 57±28                 | 1.4±1.0         | 4.2±1.5          | 1.5±1.1       | -1.9±3.1        |

<sup>+</sup>Rraup-bark; <sup>++</sup>Chutter-bark, <sup>+++</sup>Leopard, Freq = Frequency



**Figure 3.2** Scatterplots and spectrograms illustrating the acoustic structure of female (above) and male (below) alarm call elements. Scatterplots present the distribution of LDA discriminant scores, with circles indicating call elements given to leopards, triangles indicating call elements given to raptors, and squares indicating call elements given to snakes. Ellipses represent 95% confidence intervals for call elements given to each predator class that were correctly assigned to the eliciting context. Spectrograms illustrate a range of call elements given to each predator class; typical exemplars are those that were assigned by LDA to the correct context with a likelihood of >0.95, semi-typical exemplars are those that were assigned to the correct context with a likelihood of between 0.6-0.8, and atypical exemplars are those that were assigned to the correct context with a likelihood of <0.3 indicating that atypical exemplars were in fact misclassified. Spectrograms were made with a 512 FFT and a Hamming window.

Using the same acoustic parameters, LDA classified 78% of male call elements to the correct predator class. Within predator class, correct classification was high for call elements given to leopards and calls given to raptors (86% and 91% respectively), but call elements given to snakes were correctly classified in only 57% of cases (Figure 3.2b). A higher overall percentage of calls were classified correctly with the pDFA (87%), this could be explained by the call subset containing fewer calls given to snakes. Frequency range was the most influential variable in differentiating between call categories. Male call elements given to different predator classes were also structurally similar to male call clusters identified within the three-cluster solution of the k-means analysis, but whilst the group of calls given to leopards is similar in structure to earlier descriptions of a male "bark", the other two call groups identified with k-means cluster analysis are harder to place in relation to previous classifications and on the basis of their intermediate call structures they are referred to in this study as "rraup-barks" and "chutter-barks" (Table 3.1). Thus the structure of male alarm call elements was also related to predator type, again supporting Struhsaker's (1967b) and Seyfarth et al.'s (1980b) earlier observations, but with lower probability than female call elements, and bark-like call elements were produced in response to all three predator classes.

Looking at how temporal characteristics varied at the level of the whole call, we found that within female calls, predator type had a significant effect on the duration of whole calls (likelihood ratio test:  $\chi^2 = 20.10$ ,  $df=2$ ,  $P<0.001$ ), on the number of elements within a call (likelihood ratio test:  $\chi^2 = 64.14$ ,  $df=2$ ,  $P<0.001$ ), and on the mean inter-element interval (likelihood ratio test:  $\chi^2 = 33.37$ ,  $df=2$ ,  $P<0.001$ ). More specifically, leopards typically elicited calls of short duration, with few elements and relatively long inter-element intervals; raptors elicited calls also of short duration, but with an intermediate number of elements produced in very rapid succession, and snakes elicited calls of the longest duration, with the highest number of elements and intermediate inter-element intervals (Table 3.2). Within male calls, predator type did not have a significant effect on the duration of whole calls (likelihood ratio test:  $\chi^2 = 2.61$ ,  $df=2$ ,  $P=0.27$ ), but did have a significant effect on the mean inter-element interval within calls (likelihood ratio test:  $\chi^2 = 6.44$ ,  $df=2$ ,  $p<0.05$ ), and on the number of elements within a call (likelihood ratio test:  $\chi^2 = 6.30$ ,  $df=2$ ,  $p<0.05$ ). Within these different contexts, males produced more elements in response to snakes than leopards, and produced calls with shorter inter-element intervals in response to raptors than snakes or leopards (Table 3.2). Thus within male calls, predator specificity was stronger in temporal than spectral features.

In sum, LDA results indicate that despite being graded in the acoustic structure of their individual elements, female and male calls given to different predator classes are, to a high degree, structurally distinguishable. Differences in temporal features at the level of whole calls

provide additional cues by which a listener could distinguish each alarm call type from the others.

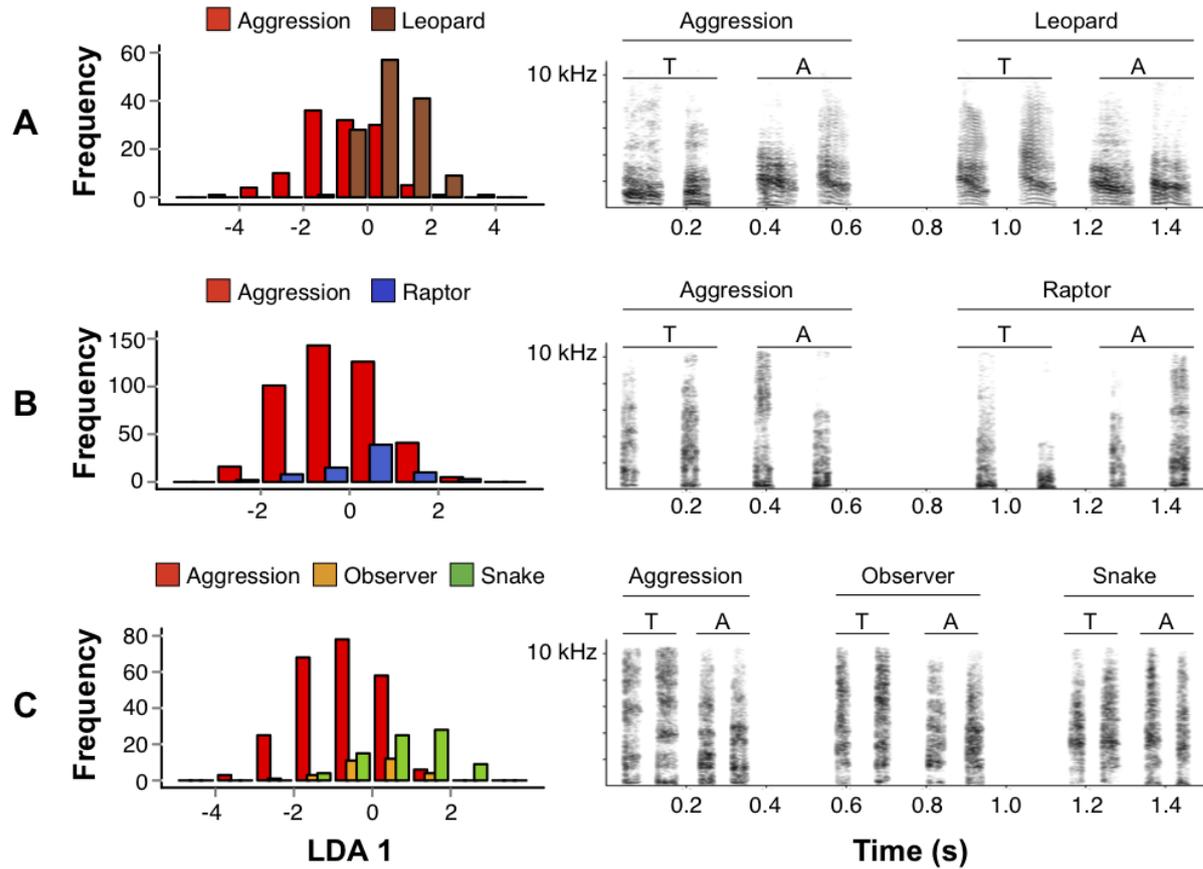
**Table 3.2** Description of temporal characteristics for female and male calls produced within predator and non-predator contexts

| Context    | Call    | Number of elements | Inter-element Interval (ms) |
|------------|---------|--------------------|-----------------------------|
| Female     |         |                    |                             |
| Leopard    | 201±217 | 1.5±0.9            | 103±105                     |
| Raptor     | 130±35  | 2.8±0.8            | 19±9                        |
| Snake      | 322±223 | 4.3±2.0            | 46.5±22                     |
| Social     | 247±199 | 4.2±2.0            | 28.0±18.2                   |
| Male       |         |                    |                             |
| E. Africa  |         |                    |                             |
| Leopard    | 628±538 | 4.4±3.0            | 93±162                      |
| Raptor     | 228±113 | 3.6±0.5            | 18±6                        |
| Snake      | 666±422 | 6.9±3.3            | 68±31                       |
| S. Africa  |         |                    |                             |
| Leopard    | 147±136 | 1.8±1.4            | 26±12                       |
| Aggression | 200±152 | 2.1±1.5            | 35±10                       |

### Predator and non-predator contexts

Based on measures of element duration, peak frequency, frequency range, PF jump and PF linear trend, LDA correctly classified the eliciting context of 54% of female chatter-like elements given to snakes, human observers and during social interactions; 60% of female rraup-like elements given to raptors and during social interactions, and 75% of male bark-like elements given in response to leopards and during aggressive interactions (Figure 3.3). LDA results were supported by pDFAs carried out on subsets of these data using the same variables, which correctly classified 60% of chatter-like elements, 57% of rraup-like elements and 74% of bark-like elements. In differentiating between elements produced in different contexts, peak frequency and frequency range were most influential in separating chatter-like elements, with those produced within social contexts possessing a lower peak frequency and larger frequency range than those produced in the presence of a snake. Frequency range was also most influential in separating rraup-like elements, again with those produced during social contexts possessing a larger frequency range than those produced in response to raptors; and peak frequency followed by PF linear trend were most influential in separating bark-like elements, with barks produced during aggression having a lower peak frequency and less steep decline of peak frequency throughout the element. In summary, when the analyses were restricted to single call elements, differences between alarm and non-alarm calls were graded rather than

discrete, and when call elements produced in non-predator contexts were included in analyses, a higher number of elements were misclassified with regards to the eliciting context.



**Figure 3.3** Histograms and spectrograms illustrating the acoustic structure of bark-like call elements (A), rraup-like call elements (B), and chirp-like call elements (C). In the left panel, histograms show the distribution of the first LDA discriminant score for bark-like (A), rraup-like (B), and chirp-like (C) call elements given in predator and non-predator contexts. In the right panel, spectrograms depict bark-like (A), rraup-like (B), and chirp-like (C) call elements given within these same predator and non-predator contexts. For each call type and context, two typical (T) exemplars and two atypical (U) exemplars are presented; typical exemplars are those that were assigned by LDA to the correct context with a likelihood of >0.9, and atypical exemplars are those that were assigned to the correct context with a likelihood of <0.45. Spectrograms were made with a 512 FFT and a Hamming window.

Concerning temporal measures, female calls produced during social interactions were longer than those produced in response to raptors (likelihood ratio test:  $\chi^2= 39.22$ ,  $df=1$ ,  $P<0.001$ ), and contained more call elements (likelihood ratio test:  $\chi^2= 39.55$ ,  $df=1$ ,  $P<0.001$ ) produced with longer inter-element intervals (likelihood ratio test:  $\chi^2= 21.52$ ,  $df=1$ ,  $P<0.001$ ).

Compared to calls elicited by snakes, calls produced during social contexts had significantly shorter inter-element intervals (likelihood ratio test:  $\chi^2= 24.41$ ,  $df=1$ ,  $P<0.001$ ), in addition they tended to be shorter but did not differ significantly with regard to duration (likelihood ratio test:  $\chi^2= 3.78$ ,  $df=1$ ,  $P=0.05$ ); nor did they differ significantly with regard to the number of elements within a whole call (likelihood ratio test:  $\chi^2= 0.19$ ,  $df=1$ ,  $P=0.66$ ). Male barks produced during aggressive interactions did not differ significantly from those produced in response to leopards with respect to call duration (likelihood ratio test:  $\chi^2= 2.51$ ,  $df=1$ ,  $P=0.11$ ), or to the number of elements (likelihood ratio test:  $\chi^2= 0.96$ ,  $df=1$ ,  $p=0.33$ ). Within calls produced during aggression, however, elements were produced with significantly longer inter-element intervals (likelihood ratio test:  $\chi^2= 5.07$ ,  $df=1$ ,  $P<0.05$ ). Descriptive statistics of differences in temporal characteristics for both female and male calls are given in Table 3.2. Thus temporal characteristics at the level of whole calls appear to be more related to predator type than spectral differences at the level of the element, especially for calls produced by females in response to raptors. Female calls given to snakes and within contexts of social interactions, and male calls produced in response to leopards and during aggressive interactions appear to differ mainly in having shorter inter-element intervals at the level of the whole call.

In sum, when elements given during alarm contexts were compared with those given in other contexts, differences were less clear than they were when alarm calls were compared with each other. Nonetheless, at the level of acoustic elements LDA correctly classified the majority of the eliciting contexts of the female and male calls we examined. Furthermore, as with our earlier comparison among the three types of alarm call, differences in temporal features at the level of whole calls provided additional cues by which a listener could distinguish between each alarm call type and the non-alarm vocalisation that it most closely resembled.

## Discussion

### Call production

When analysed at the level of single call elements, vervet alarm calls, especially those of males, are spectrally less discrete than previously supposed. In addition, when alarm call elements are compared with non-alarm call elements, they reveal a low degree of context-specificity. Despite this graded structure, vervet alarm call elements can be accurately distinguished from one another using LDA. The elements contained within male alarm calls are less distinct than those contained within female alarms, and the elements in all three alarm call types are less distinct from those contained within certain non-alarm vocalisations than they are

from each other. Thus, at the level of acoustic elements, the alarm calls of vervet monkeys show both distinctive acoustic features and some intergradation. The acoustic elements within vervet alarm calls may therefore have more in common with the graded alarm calling system of baboons (Fischer et al. 2001a) and Barbary macaques (Hammerschmidt and Fischer 1998b) than previously supposed. Interestingly, while the alarm calling systems of some arboreal cercopithecines were initially characterised as discrete (Zuberbühler 2000a; Arnold and Zuberbühler 2006a), a new study suggests that when specifically tested, one such alarm call system reveals graded exemplars as well (Keenan et al. 2013). This highlights the importance of not only looking at whether calls are structurally distinguishable but also applying cluster-based techniques to provide quantitative descriptions of call structure (Hammerschmidt and Fischer 1998b), and considering calls produced across different contexts when quantifying production specificity (Meise et al. 2011).

Importantly, vervet calls varied with regard to temporal features, specifically the number of call elements and the rate at which elements were produced. Previous studies also identified temporal characteristics as being important to distinguish between calls (Struhsaker 1967b; Owren and Bernacki 1988). Such temporal aspects are highly salient and are probably used by listeners to disambiguate calls that fall at the acoustic boundary between different call types.

What do our findings on the structure of the calls tell us about the mechanisms underlying the production of different call types? Accumulating evidence suggests that the acoustic features of calls produced by non-human primates is largely innate, although individuals have more control over temporal aspects and the usage of calls generally (Seyfarth and Cheney 2010). A more contentious issue is the question whether or not calls should be viewed as reflecting the emotional state of the animals. Without independent evidence, this is hard to decide. On the one hand, a study by Jürgens in which squirrel monkeys (*Saimiri sciureus*) could either seek or avoid electric stimulation of specific brain areas, showed that the two conditions were accompanied by specific call types that were otherwise also associated with positive or aversive stimuli (Jürgens 1979). On the other hand, there is evidence that animals are able to control the usage of calls to some degree both in captive studies (Leander et al. 1972). and in the wild (Cheney and Seyfarth 1985a), refuting the notion that calling is entirely involuntary and hardwired. Furthermore, a study of the ontogenetic development of alarm call usage showed that young vervet monkeys first over-generalised, and then slowly learned to distinguish predators from other stimuli in the environment (Seyfarth and Cheney 1980). Based on the results of our acoustic analysis alone, we are unable to resolve this issue.

In any case, the results clearly show that the structure of vervet monkey calls differs from human words (or a precursor), in that the acoustic structure is not arbitrary but rather relatively innate (Cheney and Seyfarth 1992), and shaped by selection pressures related to call function (Owren and Rendall 2001). For instance, species suffering from high predation pressure by aerial predators are more likely to evolve aerial alarms that are acoustically distinguishable from other alarm call types (Fichtel and Kappeler 2002). Because of the strong innate component, similarities in call form reflect similarities in call function. This is not a new line of argument, as Struhsaker (1967b) and Seyfarth et al. (1980b) both identified similarities between chutters produced in alarm and contexts of intra and inter-group aggression, and suggested a shared function to alert group members and solicit group defence. The loud and conspicuous bark vocalisations of adult males have been proposed as functioning to reduce the probability of a predator hunting in the vicinity again (Cheney and Seyfarth 1981), similar to the 'perception advertisement' hypothesis that alarm calls function to advertise predator presence (Zuberbühler et al. 1999b). Given the similarity of barks produced by males in response to leopards and within aggressive interactions, and the finding that male calling is related to rank (Cheney and Seyfarth 1981), it is possible that vervet alarm barks may also function as a sexual display advertising male fitness, akin to the "wahoo" of male baboons (*Papio ursinus*, Fischer et al. 2004) and the predator signals of some birds (Langmore and Mulder 1992; Ellis 2009; Greig and Pruett-Jones 2010). Further study would be needed to test whether barks function in intra and/or inter-sexual displays, and also to quantify sex differences in call structure and to investigate whether they are the result of physiological differences or sex-specific selective pressures. The relationship between the context a call is produced in and call form, function and meaning is, however, not always straightforward (Cheney and Seyfarth 1988; Hauser 1998).

Few recordings of male calls to snakes and raptors were available. This was mainly due to the difficulties of recording infrequent (crowned eagles became increasingly rare) and sometimes short calling events, and possibly also sex differences in call frequency. It is also interesting to note that the production of barks during intra- and inter-group aggression appears to be more common in South African than East African males. As aggressive encounters occur frequently between vervet groups in East Africa (Cheney and Seyfarth 1992), the apparent difference in the frequency of aggression-barks suggests a population difference in call usage. Focal data and recordings of whole calling bouts would be necessary to test this assumption, and to exclude the possibility that rarer incidences of call production have been missed. For example, Struhsaker described chirps as being produced in the later part of calling bouts to raptors as well as in response to leopards (Struhsaker 1967b), but this is not evident from our results.

In addition to functional and evolutionary aspects, it seems likely that changes in arousal level affect variation in call structure. This is in keeping with Struhsaker's (1967b) observation that temporal and structural parameters of vervet alarm calls varied throughout a calling bout in relation to proximity of the predator and, probably, caller arousal. This is also in accordance with numerous studies looking specifically at acoustic correlates of caller arousal, which show that caller arousal can affect temporal measures, frequency distributions and harmonic-to-noise ratio (reviewed in Briefer 2012).

Taken together, non-human primate vocalisations are generally more similar to the graded vocal patterns present in speech (Hockett 1960) and/or to our innate laughs, cries and screams (Fitch and Zuberbühler 2013) than the discrete phonetic units which are described as a fundamental design feature of human language (Studdert-Kennedy 2005). The allophonic aspects of phonemes within a language are, however, also graded (Peterson and Barney 1952). When viewed in this light, the production of “functionally referential” calls does not offer insights into the evolution of semantic reference at the production level in human speech (Seyfarth and Cheney 2003b).

### **Call Perception**

Vervet monkeys in the wild respond to playbacks of their alarm calls, on average, with predator-appropriate behaviours in the absence of supporting contextual cues (Seyfarth et al. 1980b), and a study with captive vervets shows that individuals are able to distinguish between female calls produced in response to snakes and raptors (Owren 1990a). Thus, it may be that vervet monkeys, when faced with ambiguous exemplars, are able to recode the acoustic variation into discrete categories, as has been found in some other primate species (Fischer 1998). In addition, it seems highly likely that vervets attend to predator-specific differences encoded in the temporal characteristics of calls, and in call combinations (Arnold and Zuberbühler 2006a). It may be the case that conspecific's difficulty in perceiving spectral differences at the level of the call element, selected for callers to produce multi-component calls (Rowe 1999). Possibly, signal receivers give more weight to call elements at the start of a calling bout (Keenan et al. 2013), but we were not able to assess call structure in relation to position within the entire calling bout, nor were we able to calculate call frequencies due to the sampling scheme.

When functionally referential calls were first proposed it was thought that, in showing parallels to the semantic properties of language, referentially specific calls that elicited appropriate responses in the absence of contextual cues were more cognitively complex than

less referentially specific calls that required contextual cues for a correct "interpretation". More recently, however, it has been proposed that we view call perception within the framework of associative learning, whereby to respond appropriately a listener must learn to integrate information from the call and the context in which it is given. (Seyfarth and Cheney 2003b). Irrespective of the mechanisms underlying call production, vocal cues may be associated with the caller's affective state, identity, behaviour, and/or an external stimulus, and the underlying cognitive mechanisms are likely no different when responding to these different cues, or indeed to other signs, such as a predator's own calls. In accordance with earlier views (Smith 1977), the "associative learning" hypothesis predicts that context-specificity is not a requisite for calls to function referentially, as contextual cues can be used to clarify call "meaning" when it is not clear from structural cues alone. The associative learning hypothesis predicts, however, that compared to the "interpretation" of highly context-specific calls possible from vocal cues alone, the integration of signal and context may, in fact, present a more complex cognitive challenge to receivers (Wheeler et al. 2011; Wheeler and Fischer 2012). Thus the associative learning framework highlights that a first step to an understanding of the cognitive mechanisms underlying call perception is not to ask "are calls functionally referential?", but rather "to what degree are calls context specific?", and to do this we need to assess a call's production specificity based on its occurrence across all contexts (Fischer 2013).

## Conclusions

Our results show that, at the level of single call elements, vervet alarm calls show considerable intergradation, both among the different alarm call types and between alarm calls and calls produced outside of predator contexts. Despite the graded spectral structure of their call elements, however, vervet alarm calls are nevertheless distinguishable with regard to predator type, both with regard to the spectral cues of these elements and with regard to the temporal cues of whole calls. Notably though, like many other animal signals, the call elements show acoustic intergradations from one call type to another and are thus fundamentally different from the discrete phonetic units that are a design feature of language. Perhaps it is now time to move beyond the metaphor of functionally referential alarm calls as precursors to symbolic communication (see Wheeler and Fischer 2012 for discussion), and focus instead on the selective pressures shaping primate vocalisations over evolutionary time.

## Methods

### Data collection

Vocalisations of East African vervets were recorded in analogue form from several free-ranging habituated groups within Amboseli National Park in Kenya, by T. Struhsaker (June 1963-May 1964) and subsequently by R. Seyfarth and D. Cheney (1977-1988). These calls were digitised post-recording at 44.1 and 22.05 kHz and 16 bit sampling depth. For a more detailed description of study subjects, study sites and recording equipment see (Struhsaker 1967b; Seyfarth et al. 1980b). South African vervet vocalisations were recorded by T. Price (Jan-June 2012), using a Marantz PMD661 solid-state recorder (44.1 kHz sampling rate; 16-bit sampling depth) connected to a Sennheiser ME66/K6 directional microphone from four free-ranging habituated groups within the Loskop Dam Nature Reserve in South Africa. Spontaneous vocalisations were recorded using focal individual and/or ad libitum sampling techniques during all contributing studies, and R. Seyfarth, D. Cheney and T. Price supplemented recordings elicited by snakes and leopards with vocalisations given in response to snake and leopard models.

### Call selection

To assess the acoustic structure of East African vervet alarm calls and the degree to which they are predator-specific, all vocalisations produced in the presence of a raptor, leopard or snake were selected. Calls given by immature and unidentified individuals were discarded to avoid age-related effects and to control for individual differences in analyses. Spectrograms of vocalisations were inspected visually using Avisoft-SASLab Pro (R. Specht Berlin, Germany, version 5.1.20); calls were selected for temporal analysis when the start and end points of all constituent call elements were clear, and exhalation call elements were selected for spectral analysis when they were undisturbed by other sounds and possessed a high signal-to-noise ratio. In some contexts individuals produced very long calling bouts consisting of hundreds of calls; to get a more balanced sample for spectral analysis we randomly selected a maximum of 16 call elements from any one calling bout. The final data set comprised a total of 498 calls for analysis of temporal characteristics and 678 call elements for analysis of spectral characteristics; Table 3.3 gives an overview of the distribution of these calls across contexts and callers.

**Table 3.3** The number of East African vocalisations used in spectral analyses to describe call structure and spectral and temporal analyses to assess context specificity within predator contexts. Call samples are described per individual and per context

| Context | Spectral Analysis of elements |                  |               |                      | Temporal Analysis of calls |               |               |                   |
|---------|-------------------------------|------------------|---------------|----------------------|----------------------------|---------------|---------------|-------------------|
|         | Call origin                   | Total # elements | Indiv- iduals | Elements/ Individual | Call origin                | Total # calls | Indiv- iduals | Calls/ Individual |
| Leopard |                               |                  |               |                      |                            |               |               |                   |
| Female  | 1                             | 186 (41)         | 17            | 1-42                 | 1                          | 171 (28)      | 17            | 1-40              |
| Male    | 1                             | 248 (53)         | 12            | 9-33                 | 1                          | 206 (33)      | 12            | 4-52              |
| Raptor  |                               |                  |               |                      |                            |               |               |                   |
| Female  | 1                             | 82               | 13            | 1-30                 | 1                          | 52            | 13            | 1-18              |
| Male    | 1&2                           | 32               | 3             | 4-15                 | 1                          | 5             | 2             | 1-4               |
| Snake   |                               |                  |               |                      |                            |               |               |                   |
| Female  | 1                             | 84 (69)          | 6             | 1-19                 | 1                          | 35 (26)       | 6             | 1-10              |
| Male    | 1&2                           | 46 (15)          | 3             | 15-16                | 1                          | 29 (21)       | 2             | 8-21              |

Call origin: 1 Calls recorded by R. Seyfarth and D. Cheney; 2 Calls recorded by T. Struhsaker

Numbers in brackets are the number of calls that were produced in response to predator models

To assess the degree to which vervet alarm calls are context-specific outside of the predator-context, we next incorporated calls produced outside of the predator context into our analyses. For females we compared the temporal characteristics of calls produced in the presence of a predator with the temporal characteristics of calls produced in response to the presence of the observer, and calls produced during social interactions. For a comparison of spectral characteristics we used the three-cluster solution of the k-means cluster analysis to select female chatter elements given to snakes and female rraup elements given to raptors. We then compared these calls to calls of a similar structure (chatter-like and rraup-like call elements) produced by adult females from the same East African population in non-predatory contexts. Whilst East African adult male vervets produce barks in contexts of aggression as well as in response to terrestrial and aerial predators (Struhsaker 1967b; Cheney and Seyfarth 1992), there were insufficient recordings of these calls from this population to allow for a structural comparison between contexts. Instead, the bark-like vocalisations of adult male South African vervets produced in contexts of aggression (within and between group chases, physical attacks and between-group confrontations) and in response to leopards and leopard models were identified both by ear and by visual inspection of spectrograms, and these calls and constituent call elements were selected for comparison. All analyses were therefore carried out comparing calls originating from the same (East or South) population. As above, for spectral analysis we randomly selected a maximum of 16 call elements from any one calling bout, and the final

dataset comprised a total of 678 calls for analysis of temporal characteristics, and 1115 call elements for analysis of spectral characteristics. Table 3.4 gives an overview of the distribution of these calls across contexts and callers.

**Table 3.4** The number of East African and South African vocalisations used in spectral and temporal analyses to assess context-specificity outside of predator contexts. Call samples are described per individual and per context

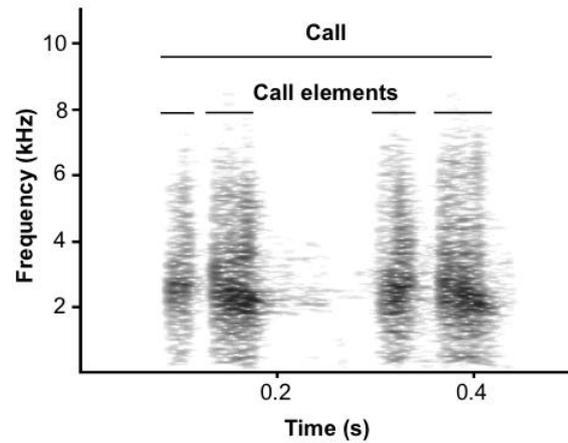
| Spectral Analysis of call elements |             |                  |             |                      | Temporal Analysis of calls |             |                  |             |                   |
|------------------------------------|-------------|------------------|-------------|----------------------|----------------------------|-------------|------------------|-------------|-------------------|
| Context                            | Call origin | Total # elements | Individuals | Elements/ Individual | Context                    | Call origin | Total # of calls | Individuals | Calls/ Individual |
| Bark-like                          |             |                  |             |                      | Male                       |             |                  |             |                   |
| Leopard                            | 1           | 137 (108)        | 10          | 10-15                | Leopard                    | 1           | 173 (146)        | 10          | 11-25             |
| Aggression                         | 1           | 119              | 9           | 3-28                 | Aggression                 | 1           | 99               | 8           | 1-25              |
| Rraup-like                         |             |                  |             |                      | Female                     |             |                  |             |                   |
| Raptor                             | 2           | 77               | 13          | 1-30                 | Raptor                     | 2           | 52               | 13          | 1-18              |
| Social                             | 2           | 432              | 27          | 1-81                 | Snake                      | 2           | 35 (26)          | 6           | 1-10              |
| Chutter-like                       |             |                  |             |                      | Social                     | 2           | 319              | 31          | 1-56              |
| Snake                              | 2           | 81 (66)          | 6           | 1-23                 |                            |             |                  |             |                   |
| Observer                           | 2           | 31               | 7           | 2-9                  |                            |             |                  |             |                   |
| Social                             | 2           | 238              | 24          | 1-44                 |                            |             |                  |             |                   |

Call Origin: 1 Vervet calls recorded from S. African population; 2 Vervet calls recorded from E. African population  
Numbers in brackets are the number of calls that were given to predator models

### Acoustic analysis

Avisoft -SASLab Pro was used to filter recordings (high-band filter set at 0.1 kHz) to remove background noise below the lowest frequency of calls. The start and end points of all call elements were labelled (as illustrated in figure 3.4) and these measures were exported for the calculation of temporal measures, which are listed in Table 3.5. Call elements were then padded with silent margins and extracted for spectral analysis. No calls contained energy above 11 kHz, and bark-like calls contained no energy above 8 kHz. To standardise the sampling frequency of calls within each data set and to optimise the frequency resolution with respect to the frequency ranges of different call types, the sampling frequency of all East African calls was therefore reduced to 22.05 kHz, and the sampling frequency of South African bark-like vocalisations was reduced to 16.0 kHz.

**Figure 3.4** Spectrogram with labels depicting a male alarm call and constituent alarm call elements produced in response to a leopard. The spectrogram was made with a 512 FFT and a Hamming window.



Call elements were then transformed in their frequency-time domain with Avisoft using a fast Fourier transformation (FFT) size of 1024 points, Hamming window and 93.75% overlap. Depending on the sampling frequency of the calls (22.1 kHz and 16.0 kHz respectively), this resulted in bandwidths of 28 or 20 Hz and a time step of 2.9 or 4 ms. The resulting frequency-time spectra were analysed with LMA (2012\_9), a custom software sound analysis tool developed by K. Hammerschmidt. LMA has been used to analyse recordings in a range of species (Naguib et al. 2001; Riede et al. 2001; Neumann et al. 2010), and the programme allows for the extraction of a wide range of robust acoustic parameters (Schrader and Hammerschmidt 1997). Because many of the calls in the data set lack harmonic structure, a range of these parameters that described energy distribution throughout the call element were selected from LMA output to characterize the general structure of these calls. Avisoft was used to extract element duration from the wav file, and to calculate Wiener entropy. A description of parameters used for analyses are given in Table 3.5.

**Table 3.5** Description of the acoustic parameters used to describe call structure

| Measurement               | Description  |
|---------------------------|--|
| Call measures             |  |
| Call duration (ms)        | Duration of call   |
| Element number            | Number of call elements within a call                            |
| Inter-element interval    | Mean inter-element interval within a call                        |
| Element measures          |  |
| Element duration (ms)     | Duration of call element   |
| Peak frequency (Hz)       | Median peak frequency  |
| Peak frequency_1-4 (Hz)   | Mean peak frequency at 1st, 2nd, 3rd and 4th temporal quartiles  |
| PF jump (Hz)              | Maximum difference between successive peak frequencies           |
| PF linear trend (-1 to 1) | Factor of linear trend of peak frequency throughout the call     |
| PF deviation (Hz)         | Mean deviation between peak frequency and linear trend           |
| First quartile_1-4 (Hz)   | Mean first frequency quartile at 1st, 2nd, 3rd and 4th temporal  |
| Second quartile_1-4 (Hz)  | Mean second frequency quartile at 1st, 2nd, 3rd and 4th temporal |
| Third quartile_1-4 (Hz)   | Mean third frequency quartile at 1st, 2nd, 3rd and 4th temporal  |
| Frequency range (Hz)      | Mean frequency range   |
| Wiener Entropy            | Mean value of noise within call. 0=pure tone, 1= random noise    |

## Statistical analysis

### A. Alarm call element structure

We used cluster analysis to group calls according to their acoustic structure, but prior to this we reduced the redundancy in the 22 acoustic parameters, which were in part highly correlated, by running a Principle Components Analysis (PCA) using the `prcomp` function of the R package "stats" (R Development Core Team 2011). In an initial PCA, element duration did not clearly load onto any one of the Factors with Eigenvalues  $\geq 1$  and was weakly correlated with the other predictor variables (largest absolute Spearman correlation = 0.29). Thus, we excluded element duration and re-ran the PCA with the remaining 21 variables. This produced a set of four factors with Eigenvalues  $\geq 1$ , which together explained 90% of the total variance. After varimax rotation, parameters related to lower frequency energy distributions loaded most strongly on the first factor, parameters related to entropy and frequency range loaded most on the second factor, parameters related to stability of the peak frequency throughout the call loaded most on the third factor, and the linear trend of the peak frequency loaded most on the fourth factor. For factor loadings of the predictor variables, Eigenvalues and percent variance explained, see Table S3.1 in supplementary.

Cluster analyses were run using the standardised values (z-scores) of element duration and the four principle components. Male and female calls were analysed separately. We used k-means clustering to compute 7 cluster solutions (2-8 clusters), and for each cluster solution we calculated the silhouette coefficient, a measure of how tightly the data is grouped within the clusters (Rousseeuw 1987). To illustrate call groupings we used Ward's hierarchical clustering (Ward 1963), whereby a dendrogram is constructed by linking individual calls to their nearest neighbour and repeating the procedure until the top hierarchic level is reached.

### **B. Context specificity**

To determine whether call elements given in response to different predator classes are spectrally distinguishable, we ran two Linear Discriminant Analyses (LDA), one for female call elements and one for male call elements, using the `lda` function of the R package "mass" (Venables and Ripley 2002) with predator class as the grouping variable. We then ran three LDAs to investigate the degree to which chatter, rraup and bark elements would be correctly classified with regards to context when compared to structurally similar call elements produced in non-alarm contexts. For all analyses element duration, peak frequency, frequency range, PF jump and PF linear trend were selected in accordance with results from the PCA and from a Spearman's correlation (these tests identified them as highly representative of overall call structure and only weakly correlated) and were entered into both LDAs. Classification scores were calculated using jack-knifed (i.e., leave one out) predictions. Because LDA classification is sensitive to unbalanced samples, we ran nested permuted discriminant function analyses (pDFA, Mundry and Sommer 2007) on a subset of the data using the same variables, which allowed us to control for caller identity.

We ran general linear mixed models (GLMMs) to test whether temporal characteristics of calls (call duration, the number of elements within a call, or within call inter-element intervals) demonstrate predator specificity. To do this we first ran GLMMs (with female and male calls analysed separately) to compare calls produced in response to different predator categories. Then, to assess specificity outside of the predator context, we ran GLMMs to compare female calls produced in response to raptor and snake presence with calls produced during social interactions, and to compare male calls produced in response to leopards with calls produced during aggressive interactions. Within all models the context of calling was included as the test factor, and caller identity was included as a random effect. Models were fitted using the function `lmer` of the R-package `lme4` (Bates et al. 2011), and likelihood ratio tests using Maximum

Likelihood were used to test the significance of the full model in relation to a null model which was comprised only of the intercept and the random effect.

Cluster analyses were run within the Matlab environment (MathWorks, Version 8.0), all other statistical analyses were carried out using R (R Development Core Team 2011). Before running the PCA, LDA, pDFA and GLMMs we ensured that assumptions were met (see S3.1- S3.4 in supplementary). Details of call subsets for pDFAs are also given in supplementary, Table S3.3a (predator context) and Table S3.3b (predator and non-predator context).

### **Abbreviations**

GLMM: general linear mixed model; LDA: linear discriminant function analysis; PCA: principle components analysis; pDFA: permuted discriminant function analysis; SD: standard deviation

### **Competing interests**

The authors declare that they have no competing interests.

### **Acknowledgements**

TTS would like to thank Peter Marler, and Stuart and Jeanne Altmann for their assistance and encouragement during the first studies of vervet alarm calls in 1963-4. TP thanks Mpumalanga parks board for permission to work at Loskop Dam Nature Reserve, and Jannie and Elizma Coetzee, Tersia de Beer and Alan Barrett for their logistical support at this field site. Oumar Ndiaye's assistance in the field was invaluable. Roger Mundry kindly provided the R script for pDFA and statistical assistance in the form of written tutorials. The comments of Kurt Hammerschmidt, Brandon Wheeler and four anonymous reviewers substantially improved the manuscript. The study was supported by funding from the Deutscher Akademischer Austauschdienst, The Leakey Foundation, The Wenner-Gren Foundation and the Federal Ministry of Education and Research (BMBF) Germany under grant number 01GQ1005C.

## Supplementary

### S3.1 Principle Components Analysis (PCA)

Prior to running the Principle Components Analysis we visually inspected histograms and scatterplots of the 22 acoustic parameters to check that there were no outliers, that their univariate distributions were approximately symmetrical, and that when there was a relation between variables that this relation was linear. Predictor variables that were not symmetrically distributed were transformed. When taking the log or square-root of the parameter did not achieve symmetrical distribution, we used the Box-Cox function (R package "mass", Ripley et al 2008) to find the optimum transformation based on maximum-likelihood methods. After transformation all acoustic parameters were standardised using a Z-transformation in order to make them more comparable in terms of their units of measure and variance.

**Table S3.1** Factor loadings of the 21 predictor variables (duration excluded), as well as Eigenvalues and percent variance explained. For each variable the largest absolute loading is indicated in bold

|                   | <b>Factor<br/>1</b> | <b>Factor<br/>2</b> | <b>Factor<br/>3</b> | <b>Factor<br/>4</b> |
|-------------------|---------------------|---------------------|---------------------|---------------------|
| Peak frequency_1  | <b>0.78</b>         | 0.21                | 0.13                | 0.33                |
| Peak frequency_2  | <b>0.82</b>         | 0.29                | 0.13                | 0.19                |
| Peak frequency_3* | <b>0.84</b>         | 0.32                | 0.11                | -0.11               |
| Peak frequency_4* | <b>0.77</b>         | 0.25                | -0.01               | -0.21               |
| First quartile_1  | <b>0.82</b>         | 0.28                | 0.09                | 0.41                |
| First quartile_2  | <b>0.88</b>         | 0.38                | 0.08                | 0.23                |
| First quartile_3  | <b>0.86</b>         | 0.48                | 0.08                | -0.09               |
| First quartile_4  | <b>0.80</b>         | 0.48                | 0.06                | -0.25               |
| Second quartile_1 | <b>0.66</b>         | 0.52                | 0.15                | 0.43                |
| Second quartile_2 | <b>0.67</b>         | 0.66                | 0.17                | 0.20                |
| Second quartile_3 | 0.64                | <b>0.71</b>         | 0.18                | -0.12               |
| Second quartile_4 | 0.61                | <b>0.66</b>         | 0.21                | -0.29               |
| Third quartile_1  | 0.44                | <b>0.71</b>         | 0.08                | 0.38                |
| Third quartile_2  | 0.39                | <b>0.85</b>         | 0.07                | 0.15                |
| Third quartile_3  | 0.37                | <b>0.89</b>         | 0.09                | -0.09               |
| Third quartile_4  | 0.40                | <b>0.80</b>         | 0.15                | -0.22               |
| Entropy           | 0.28                | <b>0.79</b>         | 0.25                | 0.17                |
| Frequency range   | 0.21                | <b>0.86</b>         | 0.27                | -0.12               |
| PF jump*          | 0.09                | 0.22                | <b>0.92</b>         | 0.06                |
| PF deviation*     | 0.11                | 0.19                | <b>0.91</b>         | 0.11                |
| PF linear trend   | -0.02               | 0.06                | -0.09               | <b>-0.59</b>        |
| <b>Eigenvalue</b> | <b>3.69</b>         | <b>1.45</b>         | <b>1.4</b>          | <b>1.12</b>         |
| <b>% variance</b> | <b>65</b>           | <b>10</b>           | <b>9</b>            | <b>6</b>            |

\*Transformed prior to entry into the PCA. Peak frequency\_3 and peak frequency\_4 transformed to  $\wedge^{0.7}$  and  $\wedge^{0.5}$  respectively, with duration (in initial PCA) and PF deviation transformed to  $\wedge^{0.1}$ , and PF jump transformed to  $\wedge^{0.2}$ .

### S3.2 Linear Discriminant Function Analysis (LDA)

Before running LDAs to assess how call structure varied between i) alarm calls, and ii) alarm and non-alarm calls, variables were transformed using the same method as described for the PCA in order to meet the assumption of symmetric univariate distribution (see Table S3.2 below for which transformation was applied to each parameter). All acoustic parameters were then standardised using a Z-transformation.

To check the multivariate assumptions of LDA we ran a MANOVA, with context as the independent variable and the Z-scores of duration and the four principle component factors as dependant variables. We visually inspected plots of residuals to assess homogeneity and normality, and plotted mahalanobis distance to assess multivariate normality. We identified two outliers amongst mahalanobis distance values of male calls when assessing the context specificity of calls given in contexts of predator presence analysis of predator presence. We excluded these data points from analyses, following which data indicated no obvious deviations from these assumptions. Whether analyses included or excluded these outliers did not affect the significance of results, and only analyses without these outlying cases are reported.

**Table S3.2** Transformations applied to the raw data of acoustic parameters prior to statistical analysis.

| Analysis                | Subset       | Parameters transformed   |
|-------------------------|--------------|--|
| Predator                | Female       | Duration (log), PF jump ( $\wedge 0.2$ ), PF_linear trend (+9 then square root)        |
|                         | Male         | Duration (square root), PF jump (log)  |
| Predator & non-predator | Chutter-like | Peak frequency ( $\wedge 0.4$ ), Frequency range ( $\wedge 2$ ), PF jump (square root) |
|                         | Rraup-like   | Duration ( $\wedge -0.5$ ), PF (square root), PF jump (square root)                    |
|                         | Bark-like    | PF jump (log)  |

### S3.3 Permuted Discriminant Function Analysis (pDFA)

In adherence to the requirements of nested pDFAs, calls from a single individual were selected only from one level of the test factor, and only if four or more calls were available from within this single level. See Table S3.3a and S3.3b below for details of the call samples incorporated within data subsets.

**Table S3.3a** The number of vocalisations used in pDFA to assess context specificity within the predator context; described per individual and per predator context

| Sex    | Context | Total calls | Individuals | Calls/Individual |
|--------|---------|-------------|-------------|------------------|
| Female | Leopard | 72          | 8           | 4-19             |
|        | Raptor  | 69          | 8           | 4-30             |
|        | Snake   | 83          | 5           | 10-23            |
| Male   | Leopard | 189         | 10          | 9-33             |
|        | Raptor  | 32          | 3           | 4-15             |
|        | Snake   | 30          | 2           | 15               |

**Table S3.3b** The number of vocalisations used in pDFA to assess context specificity within predator and non-predator contexts; described per individual and per context

| Call type    | Context    | Total calls | Individuals | Calls/Individual |
|--------------|------------|-------------|-------------|------------------|
| Bark-like    | Leopard    | 107         | 8           | 10-15            |
|              | Aggression | 116         | 8           | 4-28             |
| Raup-like    | Raptor     | 68          | 8           | 4-30             |
|              | Aggression | 180         | 13          | 4-48             |
| Chutter-like | Snake      | 70          | 4           | 12-23            |
|              | Observer   | 21          | 3           | 6-9              |
|              | Aggression | 199         | 15          | 4-44             |

**S3.4 General Linear Mixed Model (GLMM)**

We visually assessed histograms to check for symmetric distribution of temporal response variables. When variables did not meet this assumption we either square-root or log transformed them to achieve an approximately symmetrical distribution. Following this we inspected qqplots and residuals plotted against fitted values for all GLMMs to check that models fulfilled assumptions of normally distributed and homogenous residuals. In two tests, element number did not meet these assumptions, and in these cases we ran the GLMMs with Poisson rather than Gaussian error structure after first testing that data met the assumption of not being over-dispersed. Details of data transformations and the family of GLMM used for each test are given below in Table S3.4

**Table S3.4** Transformations applied to the raw data of temporal measures prior to statistical analysis, and family of GLMM applied

| GLMM         | Duration       | Element number | Inter-interval interval |
|--------------|----------------|----------------|-------------------------|
| Predator     |                |                |                         |
| Female       | Log; Gaussian  | None; Poisson  | Log; Gaussian           |
| Male         | Sqrt; Gaussian | Log; Gaussian  | Sqrt; Gaussian          |
| Non-Predator |                |                |                         |
| Female       | Log; Gaussian  | Log; Gaussian  | Log; Gaussian           |
| Male         | Log; Gaussian  | None; Poisson  | None; Gaussian          |

Sqrt = square-root



## Chapter 4

### **Meaning attribution in the West African green monkey: Influence of call type and context**

Tabitha Price<sup>1,2</sup>, Julia Fischer<sup>1,2</sup>

<sup>1</sup>Cognitive Ethology Lab, German Primate Center, Germany

<sup>2</sup>Courant Research Centre for the Evolution of Social Behaviour, Georg August University of  
Göttingen, Germany

Animal Cognition (2013)

10.1007/s10071-013-0660-9

### **Abstract**

The search for the evolutionary roots of human language has fuelled much research into the cognitive mechanisms underlying communication in non-human animals. One core issue has been whether the context-specific calls of non-human animals are meaningful, with call meaning inferred from recipients' responses in the absence of supporting contextual cues. This direct inference may well offer an oversimplified view of how vocalisations are perceived, however, as responses under natural conditions are likely guided by contextual cues as well as by the signal. In this study, we investigate how the anti-predator responses of green monkeys, *Chlorocebus sabaues*, are affected by alarm call structure and by context. We first simulated the presence of leopards and snakes to elicit alarm vocalisations and to identify predator-typical response behaviours. In both contexts, the monkeys produced chirp calls that revealed only graded variation in relation to predator type. We then carried out playback experiments to explore whether green monkeys would respond with predator-typical behaviour to leopard and snake chirps, and whether contextual cues, in the form of pre-exposure to a leopard or snake model, would modify these responses. Irrespective of context, subjects were more likely to respond to leopard chirps with a leopard-typical response. Predator priming did not have a significant effect on the type of response, but, together with call type, did affect response duration. This suggests that the immediate attribution of meaning was influenced by acoustic cues, while receiver's prior knowledge was incorporated to guide subsequent behaviour.

## Introduction

What do the vocalisations of animals mean? This question is central to the debate regarding the similarities and differences between non-human animal (hereafter animal) communication and human language, and consequently, language evolution. The finding that vervet monkeys (*Chlorocebus pygerythrus*) produce predator-specific alarm calls that elicit appropriate response behaviours even in the absence of contextual cues led initially to claims that these calls possessed semantic properties (Seyfarth et al. 1980a). The general consensus that, within animal communication, signallers and receivers do not share a representational state and are not motivated to communicate as a result of attributing mental states to one another (Seyfarth and Cheney 1993; Rendall et al. 2009) implies, however, that animal vocalisations are not meaningful in the linguistic sense of the word (Cheney and Seyfarth 1992; Rendall et al. 2009; Scarantino 2010).

Over the last 20 years, signals that are elicited only by stimuli belonging to a common category (i.e. are context specific) and that cause signal receivers to respond with stimulus-appropriate behaviours even in the absence of contextual cues have been termed “functionally referential” (Marler et al. 1992; Macedonia and Evans 1993). This terminology was meant to emphasize that such signals are “not exactly like human words, but rather appear to function in the same way” (Hauser 1997 p. 509). Numerous studies indicate that receiver responses cannot be explained only in terms of unconditioned reactions to the acoustic properties of a call (reviewed in Seyfarth et al. 2010), or by perceptual similarities between the call and the stimulus (Zuberbühler et al. 1999a). Instead, across a broad array of taxa, signal receivers respond to calls as if they had learnt to associate them with a specific predator class (Manser et al. 2001; Gill and Sealy 2004; Kirchhof and Hammerschmidt 2006), degree of risk (Furrer and Manser 2009), food (Evans and Evans 2007), social situation (Faragó et al. 2010), and/or individual (Cheney and Seyfarth 1982a; Vignal et al. 2008). It is worth noting, however, that this is not a universal property of calls. The alarm calls of American red squirrels, for example, demonstrate low predator specificity (Digweed and Rendall 2009), and the recruitment calls of the banded mongoose convey information about the risk posed by a stimulus rather than stimulus type (Furrer and Manser 2009). In addition, while the vocalisations of many species are structurally discrete, this is not a pre-requisite for functional reference; context specific calls that differ along a graded continuum may also elicit appropriate responses from signal receivers in the absence of supporting contextual cues (Fischer 1998), although this ability may require a degree of learning (Fischer et al. 2000).

The above description of receivers associating calls with referents is in line with insights from learning theory and more specifically Pavlovian conditioning (reviewed in Rescorla 1988), whereby functionally referential alarm calls can be classified as a conditioned stimulus (Seyfarth and Cheney 2003b) with an indexical relationship between the call and referent (reviewed in Wheeler and Fischer 2012). But whilst laboratory experiments within the framework of learning theory have shown effects of context specificity on the initial formation, extinction and renewal of conditioned responses in humans and other animals (Bouton et al. 2006; Huff et al. 2011), and identified neurological mechanisms underlying these effects (Hobin et al. 2003), the current definition of functional reference requires the attribution of meaning in the absence of relevant contextual cues. An alternative proposal in keeping with the influence of context on meaning attribution is that context specificity is not a requirement for calls to function referentially, only that the less referentially specific a call is, the more important contextual cues will be for an accurate attribution of meaning (Scarantino in press; Wheeler and Fischer 2012). In this study we therefore use meaning to refer to what the signal receiver infers from a signal, for example the presence of an external stimulus or the subsequent behaviour of the signaller.

Studies of animal communication have shown that the response behaviours of signal receivers are, in some cases, modified by contextual cues, including the signal receiver's prior experience (Zuberbühler 2000c; Engh et al. 2006; Akçay et al. 2009; Arnold and Zuberbühler 2013), and contextual cues at the time of hearing a call (Rendall et al. 1999; Wheeler and Hammerschmidt 2013), which may include the presence or absence of additional signals (e.g. multimodal signals; reviewed in Partan and Marler 1999). But despite this, and the fact that the role of context on call perception presents a possible parallel with pragmatics in human language (Scott-Phillips 2009; Wheeler et al. 2011), we know little about how context specificity and structure (discrete versus graded) of a call affect the degree to which contextual cues are incorporated.

More than forty years have gone by since Struhsaker (1967b) described the vervet monkey's predator specific alarm calls, and they remain the classic example of functional reference within the animal kingdom. However, a relatively high number of individuals did not respond appropriately to alarm calls when they were broadcast in the absence of supporting contextual cues (Seyfarth et al. 1980b); and chirps are described as being produced in response to both avian and major terrestrial predators (Struhsaker 1967b). Taken together it seems likely that both context and call structure contribute to the attribution of call meaning by conspecifics.

Like adult female vervets, adult female green monkeys (*C. sabaesus*) produce chirp calls in response to more than one predator class. The green monkey is a close relative of the vervet, and they were previously classified as conspecifics (Napier 1981). We here follow the taxonomy

of Groves (2001), however, which places the green monkey as a closely related congener to the vervet. In the case of green monkeys, females produce chirp calls to both snake and leopard models (hereafter referred to as "snake chirps" and "leopard chirps"), and these calls sound acoustically similar to one another. In this study, we first investigated predator-specific behaviours in the green monkeys and analysed the acoustic structure of snake and leopard alarm chirps. We then performed experiments in which subjects were exposed to a predator model (leopard or snake) before playing back a leopard or snake chirp. If chirp calls given to leopards and snakes are strongly referential, they should elicit predator-typical avoidance behaviours irrespective of supporting or conflicting contextual cues. If, however, context also plays a role in how conspecifics' attribute meaning to these calls, then priming with a corresponding predator model (e.g. priming with a leopard model prior to playing a leopard chirp) should increase the occurrence of predator typical responses relative to responses elicited by the calls alone, whilst priming with a conflicting predator model (i.e. priming with a snake model prior to playing a leopard chirp) should have the opposite effect.

### **Study site and subjects**

The study was conducted over two field seasons (January-June 2010 and December 2010-June 2011) within Niokolo Koba National Park in southeast Senegal (13°01'34"N, 13°17'41"W), an area encompassing 913,000 ha of predominantly Sudano-Guinean savannah interspersed with woodland and gallery forest (Frederiksen and Lawesson 1992). Green monkeys are found throughout the park, living in species typical multi-male multi-female groups (Dunbar 1974). Data were collected in the vicinity of the Simenti Centre de Recherche de Primatologie (CRP Simenti) from four groups of free-ranging green monkeys ("Simenti" 16-21 individuals; "Mare" 12-18 individuals; "Lions" 19-26 individuals; "Niokolo" 27-32 individuals; ranges reflect changes in group size over the duration of the study period). Study subjects were habituated adult males and females that were recognised individually from natural markings on the face and body. Pythons, venomous snakes and leopards were all observed in the vicinity of the field site over the course of the study.

## **A. Behavioural response to terrestrial predators**

### **Experimental protocol**

Vervet monkeys tend to respond to snakes by looking down and standing bipedally, and to leopards by climbing up into trees (Cheney and Seyfarth 1992). To test whether green monkeys respond to these terrestrial predators with these same predator typical behaviours, we simulated the presence of snakes and leopards and video-taped their behavioural response. For details of predator simulations and modes of presentation see supplementary S4.1. Subjects were provisioned with peanuts prior to model presentation to position individuals on the ground and to ensure that subject behaviour (stationary feeding) was consistent in the time period preceding all playbacks. Experiments were discarded if the subject moved out of sight within the first 10 seconds of the experiment (5 cases), if the subject responded to a different stimulus prior to model presentation (3 cases) or if there were technical problems with the equipment (1 case), resulting in a total of 17 leopard model (adult female  $n=8$ , adult male  $n=9$ ) and 19 snake model (adult female  $n=9$ , adult male  $n=10$ ) experiments for analysis.

### **Behavioural analysis**

Behavioural responses of subjects were filmed using a Sony Handycam (DCR-HC90E), and videos were imported into Adobe Premiere Pro CS4 with a time resolution of 25 frames/second. Frame-by-frame analysis set at five-frame jumps was used to score the subject's behaviour as one of four mutually exclusive categories (rest, bipedal, terrestrial displacement or arboreal displacement) at 0.2 s intervals for a period of 10 s, starting with the subject's first response to the predator model. We had initially planned to include looking direction as a behavioural measure, but poor visibility made it impossible to score this reliably from the videos. Maximum height of the subject within 30 s of viewing the model was recorded as 0m, >0m but <2m or >2m. Because video encoding is susceptible to observer-bias, all videos were reanalysed by a second condition-naive observer. Intra-class correlation coefficient (ICC) was 0.986, indicating a high level of inter-observer reliability.

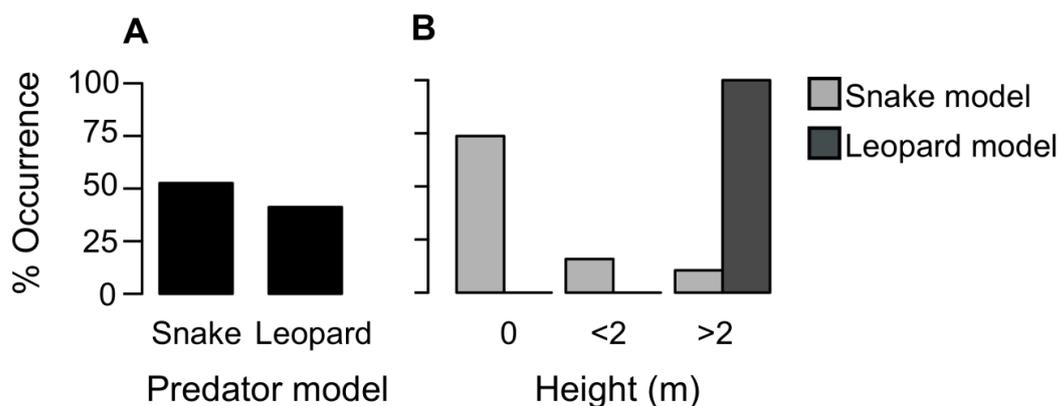
### **Statistical analysis**

We used a Generalised Linear Mixed Model (GLMM) with binomial error structure and logit link function to test whether snake models were more likely than leopard models to elicit bipedal behaviour, with bipedal behaviour scored as absent or present. A second GLMM with Poisson error structure and a log link function was run to test whether leopard models would

cause subjects to climb into a tree more often than snake models, with response behaviour scored as one of the three height categories described above. Both GLMMs were run with the type of predator model (snake or leopard) as the fixed effect and subject identity included as a random effect using the function `lmer` of the `lme4` Package (Bates et al. 2011). We used a likelihood ratio test (anova using "Chisq" argument) to compare the full models with a null model (comprising only the intercept and the random effect) in order to calculate the overall effect of the predator model. All models were fitted in R (R Development Core Team 2011).

## Results and Discussion

There was no significant difference in the bipedal behaviour of subjects following the presentation of snake and leopard models (likelihood ratio test:  $\chi^2=0.47$ ,  $df=1$ ,  $P=0.491$ ; Figure 4.1a). Like vervet monkeys, green monkeys do sometimes respond to snakes by standing bipedally, but since they also responded to leopard models with bipedal behaviour, this did not constitute a predator specific response. Whilst vervet monkeys were described as responding with bipedal behaviour to snakes, they responded to playbacks of alarm calls given to both snakes and to leopards with bipedal behaviour (Seyfarth et al. 1980b). For vervets and green monkeys, bipedalism may therefore function not only as a mobbing behaviour but also as a form of unspecific vigilance. As we were not able to assess gaze direction, we cannot discount that bipedalism for the purpose of either scanning the ground for snakes, or scanning the horizon for cats, could constitute a predator-specific response. In consequence, from the results described in this section it is not possible to identify a snake specific behavioural response with which the referential specificity of snake chirps, with and without contextual cues, could be tested.



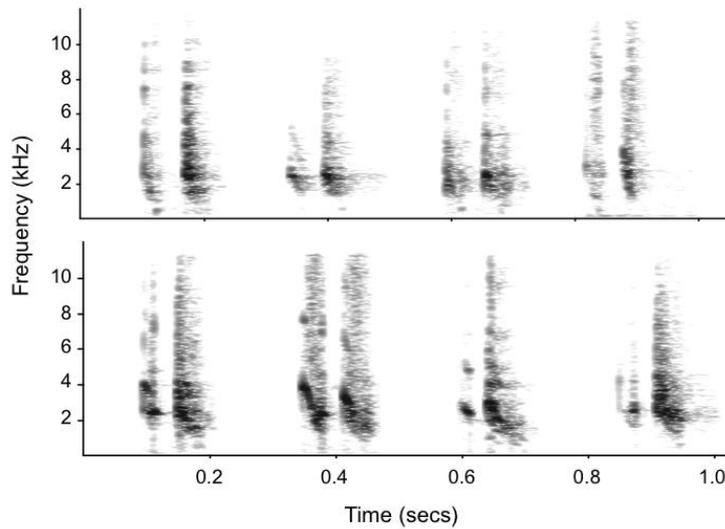
**Figure 4.1** Bar plots illustrating subjects' behavioural responses to snake ( $n = 19$ ) and leopard ( $n = 17$ ) models. a The percentage of individuals that stood bipedally within 10 s of seeing a predator model. b The maximum height of subjects within 30 s of seeing a predator model

Green monkeys, like vervets, were more likely to climb into a tree in response to leopard than snake models (likelihood ratio test:  $\chi^2 = 22.49$ ,  $df=1$ ,  $P < 0.001$ , Figure 4.1b). In particular, whilst snake models occasionally prompted subjects to jump into trees at  $< 2\text{m}$ , leopard models always resulted in subjects climbing higher ( $> 2\text{m}$ ) into a tree. This can be explained as an adaptive response, whereby green monkeys, like vervets, are likely safest from leopards high up in the trees (Cheney and Seyfarth 1992). Thus it would seem that climbing  $> 2\text{m}$  into a tree is a more leopard-specific response than simply climbing into a tree.

## **B. Chirp playback stimuli**

### **Playback stimuli**

Alarm chirps used as playback stimuli were elicited by the presentation of leopard and snake models. Calls were recorded from adult females and juveniles from all four study groups using a Marantz PMD661 solid-state recorder (44.1 kHz sampling rate; 16-bit sampling depth) connected to a Sennheiser ME66/K6 directional microphone. Vocal recordings were transferred to a PC, and Avisoft-SASLab Pro (R. Specht Berlin, Germany, version 5.1.20) was used to check recording quality, filter recordings to remove background noise below 0.1 kHz, and to prepare the playback stimuli. Each playback sequence was constructed from chirps produced during a single calling bout, although not always in their natural order, as it was sometimes necessary to replace low quality chirps with higher quality exemplars produced later in the calling bout. A total of ten pairs of playback sequences were compiled, whereby each pair consisted of a sequence of chirps given to a leopard model, and a sequence of chirps given to a snake model. The number of chirps, inter-call durations, and sequence duration was consistent between paired sequences, all call sequences were normalised to the same maximum volume, and inter-call durations were additionally controlled to fall within the range of naturally emitted calls. When possible, the same individual produced both call sequences within a pair, and at all times call sequences within a pair were produced by a caller from the same social group. With one exception, all leopard chirp and all snake chirp playback stimuli were taken from the calling bouts of different individuals, and in this exception, different calls from the same individual were used to construct two playback sequences. Calls of non-predatory birds were recorded locally and modified to be of a similar length and volume to chirp sequences for use as control stimuli. To avoid pseudo-replication, a different playback sequence was used for each playback experiment. Spectrograms illustrating snake and leopard chirps are shown in Figure 4.2.



**Figure 4.2** Spectrograms of paired chirp calls given to model snakes (top row) and model leopards (bottom row). For each context, the calls of four individuals are presented. The same individuals do not contribute calls for both contexts. Spectrograms were made with a 512 FFT and a Hamming window

### Acoustic analysis

To assess the acoustic structure of chirp calls used as playback stimuli ( $N=124$ ), Avisoft-SASLab Pro was used to add silent margins and reduce the sampling frequency of single call units to 22.05 kHz. Call units were then transformed in their frequency-time domain using a fast Fourier transformation (FFT) size of 1024 points, Hamming window and 93.75% overlap. The resulting frequency-time spectra were analysed with LMA (K. Hammerschmidt, version 2012\_9), a custom software sound analysis tool (Schrader and Hammerschmidt 1997). Using Avisoft, duration was extracted from the wav file and Wiener entropy was calculated; LMA was used to calculate robust acoustic parameters describing energy distribution throughout the call unit. A description of parameters used for analyses are given in Table 4.1.

**Table 4.1** Description of the acoustic parameters used to describe chirp call structure

| Measurement                   | Description  |
|-------------------------------|--|
| Duration (ms)                 | Duration of call unit  |
| Peak frequency_1-4 (Hz)       | Mean peak frequency at 1st, 2nd, 3rd and 4th temporal quartiles            |
| First quartile_1-4 (Hz)       | Mean first frequency quartile at 1st, 2nd, 3rd and 4th temporal quartiles  |
| Second quartile_1-4 (Hz)      | Mean second frequency quartile at 1st, 2nd, 3rd and 4th temporal quartiles |
| Third quartile_1-4 (Hz)       | Mean third frequency quartile at 1st, 2nd, 3rd and 4th temporal quartiles  |
| Wiener entropy                | Mean value of noise within call. 0=pure tone, 1= random noise              |
| Frequency range (Hz)          | Mean frequency range   |
| Jump (Hz)                     | Maximum difference between successive peak frequencies                     |
| Peak frequency deviation (Hz) | Mean deviation between peak frequency and linear trend                     |
| Linear trend                  | Factor of linear trend of peak frequency                                   |

### Statistical analysis

To avoid entering correlated parameters into the Discriminant Function Analysis, a stepwise variable selection with leave-one-out cross-validation (stepclass function of the R package “klaR”, Weihs et al. 2005) was used to first identify an optimum subset of variables for classification. Acoustic parameters were transformed when necessary to meet test assumptions (see supplementary S4.2) and then entered into the stepwise classification, with predator type set as the grouping variable. Following this, the selected variables were entered (post z-transformation) into a Linear Discriminant Function Analysis (LDA) using the lda function of the R package “mass” (Venables and Ripley 2002), with predator type again set as the grouping variable. A leave-one-out procedure was used to calculate the percentage of calls correctly classified, and a subset of the data ( $N=93$ ) was entered into a nested Permuted Discriminant Function Analysis (pDFA, Mundry and Sommer 2007) to re-calculate classification scores whilst controlling for caller identity.

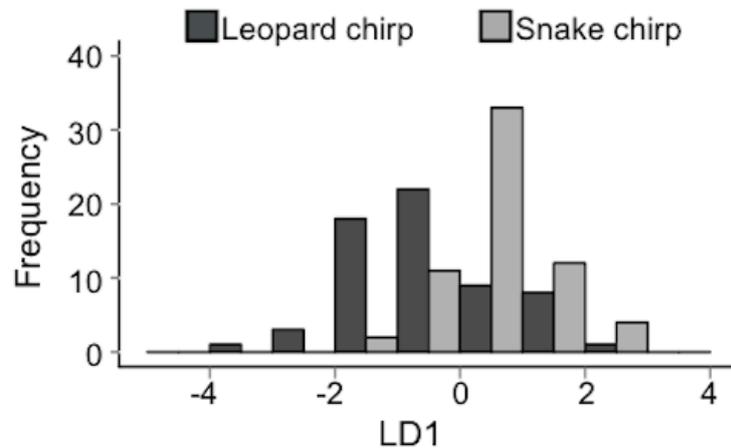
### Results and Discussion

Stepwise variable selection identified duration and peak frequency<sub>1</sub> as the most important variables for differentiating between chirps produced in response to different predator types. Based on differences in these two variables, LDA (with leave-one out validation) correctly identified leopard and snake chirps in 75% of cases. A similar result was found using a pDFA on a subset of the calls in order to control for caller identity, with 72% of calls correctly classified. On the basis of the LDA classification, chirp calls were correctly assigned to the predator type eliciting calling more often than would be expected by chance (Binomial test, chirps  $N=124$ ,  $P<0.05$ ), and each snake playback stimulus (with one exception) had a higher mean discriminant score than the leopard playback stimulus with which it was paired. The relatively high number of calls that were incorrectly classified, however, supports the acoustic impression that structural differences between leopard and snake chirps are graded rather than discrete in nature (Figure 4.3), suggesting that, for many calls, receivers would be unable to determine whether the signal was indicative of the presence of either a leopard or a snake.

Duration contributed most to distinguishing between leopard and snake chirps, followed by peak frequency<sub>1</sub>, with leopard chirps being longer than snake chirps and demonstrating a higher early peak frequency. Studies in a broad array of species suggest that as callers experience an increase in arousal, their vocalisations become longer and higher in frequency (reviewed in Briefer 2012). In line with these findings, the structural differences identified in this

study between snake and leopard chirps could be attributed to callers being more aroused in the presence of a leopard than a snake.

**Figure 4.3** Histogram showing the distribution of the first linear discriminant scores for chirp calls given in response to leopard ( $n = 62$ ) and snake ( $n = 62$ ) models



This analysis does not allow for conclusions to be drawn about the probability of chirps being produced in the presence of a snake or leopard, or whether chirps are also produced in non-predator contexts. Results do suggest, however, that the chirp is similar to the graded alarm calls of Barbary macaques (Fischer et al. 1995) and chacma baboons (Fischer et al. 2001a). Given that these two species differ in how they perceive the graded variation in their calls (Fischer 1998; Fischer et al. 2001b), the graded structure of chirps presents an opportunity to further our more general understanding of how signal receivers respond appropriately to acoustically similar alarm calls produced in situations requiring incompatible escape behaviours.

## C. Playback experiments

### Experimental protocol

To test whether green monkey leopard chirps function referentially in that they (and not snake chirps) elicit leopard appropriate responses, and to investigate whether these responses are additionally influenced by supporting and conflicting contextual cues (presence of a leopard or snake simulated by a predator model), we used a within-subjects prime and probe playback design, with each of ten experimental subjects experiencing three un-primed and four primed conditions (Table 4.2). We balanced the order in which playback stimuli were presented and included call sequence as a fixed variable within statistical analyses.

**Table 4.2** Description of the seven experimental conditions making up the prime-and probe experimental design

| Condition | Prime         | Probe         |
|-----------|---------------|---------------|
| 1         | None          | Control       |
| 2         | None          | Leopard chirp |
| 3         | None          | Snake chirp   |
| 4         | Leopard model | Leopard chirp |
| 5         | Leopard model | Snake chirp   |
| 6         | Snake model   | Leopard chirp |
| 7         | Snake model   | Snake chirp   |

Subjects were provisioned prior to playback experiments to position them on the ground at 8-15 m from a playback speaker that was hidden from sight behind a natural obstacle at a height of 1-1.5 m. Playback stimuli were broadcast using a Marantz PMD-661 solid-state recorder connected to a loudspeaker (David Active, Visonik, Berlin), with maximum amplitude set within the range of natural calling behaviour (60-85dB at 10m from source, measured using a Voltcraft 322 sound level meter). For primed conditions, predator models were presented using the same protocol as described in supplementary 4.1. When all alarm calling had stopped, a stop clock was started and a playback experiment was carried out as soon as possible within a one-hour time window. Subjects were played the chirp calls of an individual from the same group as themselves, and playbacks were carried out only when this individual was out of view. Experiments carried out on each subject were separated by  $\geq 7$  days, and a maximum of 3 playbacks (including a single leopard prime and/or a single snake prime) were carried out each week within a single group. Experiments were discarded if the wrong subject was filmed (2 cases), if the subject responded to a different stimulus prior to model presentation (1 case) or if there were technical problems with the equipment (4 cases).

### **Behavioural analysis**

Behavioural responses of subjects were filmed, videos were imported into Adobe Premiere CS4, and frame-by-frame analysis was used to score the subject's behaviour as rest, bipedal, terrestrial displacement or arboreal displacement at 0.2 second intervals as described in the section of this manuscript looking at behavioural responses to predators. Video analysis started with the onset of the playback stimuli, and continued for a period of 30 seconds. At the end of these 30 seconds, maximum height of the subject was recorded as 0m, <2m or >2m. All subjects that responded with arboreal displacement did so immediately following initiation of the

playback (subject in tree within  $1.42 \pm 0.75$  seconds), and the time a subject spent arboreal was also measured, from when the subject entered a tree until the time when the subject returned to the ground. All videos were re-assessed by two condition-naive observers, and there was a high level of inter-observer reliability (intra-class correlation coefficient = 0.996).

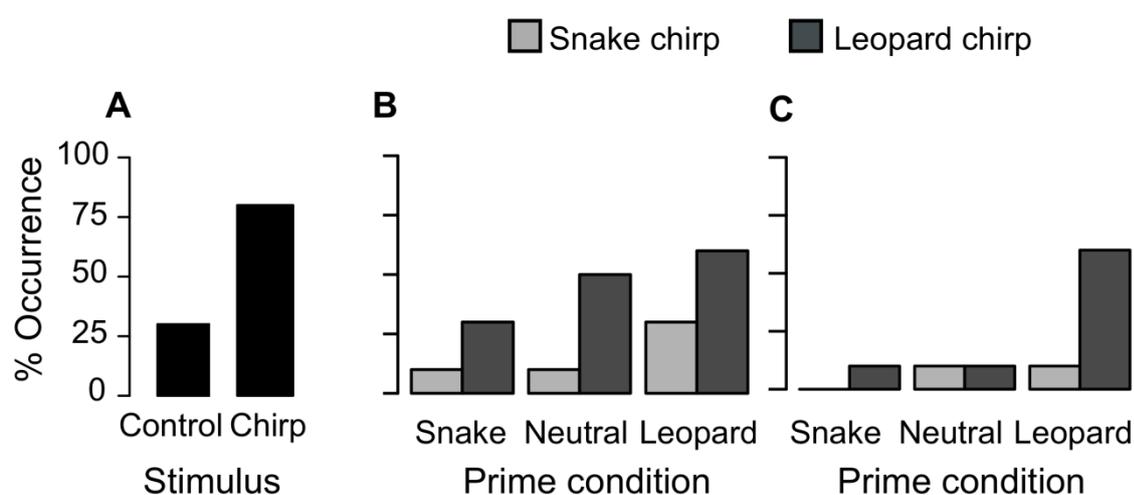
### **Statistical analysis**

To ensure that subjects' were responding to playbacks as a result of the call type and not the playback procedure itself, we ran a GLMM with binomial error structure and logit link function to model the likelihood that a subject would respond to a test versus control stimuli with any of bipedalism, terrestrial or arboreal displacement. Stimulus type (un-primed chirp or birdsong) was entered as the test predictor, playback order was entered as the control predictor (both as fixed effects), and subject identity was entered as a random effect. A likelihood ratio test was used to compare the full model with a null model, which retained all variables except stimulus type.

To assess the effect of call type and context on whether subjects would respond with a leopard-typical response, we ran a second GLMM with binomial error structure and logit link function to test differences in subjects' propensity to climb >2m into a tree. A third GLMM was run to assess whether call type or context would affect the amount of time individuals spent in a tree immediately after a playback experiment. We initially transformed the time that individuals spent in a tree into ordinal data, and used a Poisson error structure to model differences, but because the data were still overdispersed, we subsequently used a binomial error structure and logit link function to look at whether individuals stayed arboreal for longer than 200 s or not. We included call type (leopard or snake chirp), context (no prime, snake prime or leopard prime) and the interaction between the two as test predictors (fixed effects). Playback order was included as a control predictor (fixed effect) and subject identity and caller identity were included as random effects. We established the significance of the full model as compared to the null model (lacking all test predictors), and the full model as compared to reduced models (lacking the interaction and/or lacking the interaction and a test predictor) using a likelihood ratio test. Variance Inflation Factors were derived using the `vif` function of the R-package `car` (Fox and Weisberg 2011), and indicated that collinearity was not an issue. All models were fitted in R using the function `lmer` of the R-Package `lme4`.

## Results and Discussion

Subjects were significantly more likely to respond to playbacks of chirps than playbacks of bird calls (likelihood ratio test:  $\chi^2=7.76$ ,  $df=1$ ,  $P<0.01$ , Figure 4.4a). Behavioural responses to playbacks of chirp calls are thus due to signallers responding to the acoustic features of chirp calls, and not to some aspect of the playback process. In tests of whether subjects climbed to  $>2\text{m}$  in a tree, the full model explained significantly more variation than the null model ( $\chi^2=12.21$ ,  $df=5$ ,  $P<0.05$ ), although only the effect of call type (with subjects climbing higher into a tree after hearing leopard than snake chirps;  $\chi^2=8.17$ ,  $df=1$ ,  $P<0.01$ ), but not prime stimulus ( $\chi^2=3.28$ ,  $df=2$ ,  $P=0.19$ ), was significant (Figure 4.4b). In tests of the amount of time subjects spent in a tree immediately subsequent to playback experiments, the full model also explained significantly more of the variation than the null model ( $\chi^2=14.44$ ,  $df=5$ ,  $P<0.05$ ), but this time this effect was due not only to a significant effect of call type ( $\chi^2=4.90$ ,  $df=1$ ,  $P<0.05$ ), with subjects spending longer in a tree after hearing leopard than snake chirps, but also to a significant effect of context ( $\chi^2=7.41$ ,  $df=2$ ,  $P<0.05$ ) with subjects spending more time in a tree after being primed with a leopard model, Figure 4.4c).



**Figure 4.4** Bar graphs illustrating the percentage of trials in which subjects a responded to playbacks of control ( $n = 10$ ) and chirp ( $n = 20$ ) stimuli; b climbed to  $> 2\text{ m}$  within 30 s of hearing the playback stimuli, and c stayed  $> 200\text{ s}$  in a tree subsequent to hearing the playback stimuli. For playback experiments,  $n = 10$  for all conditions

That subjects were more likely to climb >2m into a tree in response to leopard chirps than to snake chirps, irrespective of contextual cues, suggests that green monkeys discriminated between graded variants of this alarm call and responded more often to leopard chirps as if a leopard were present. Given that the number of chirp units and the inter-unit duration between chirp units was kept constant between paired playbacks, this ability to discriminate between calls is apparently due to differences in call structure. At the same time, the structural similarity of the two chirp types suggests that differences in behavioural response are unlikely to be explained exclusively by unconditioned reactions to the acoustic properties of a call. Instead, it is likely that subjects' responses are the result of a learnt association, which could be underpinned by subjects associating the call with the external referent (leopard), or with the emotional response experienced by listeners via "affect conditioning ("affect-conditioning", Owren and Rendall 1997). At the ultimate level, it has been claimed that selection pressures act on receiver's "data-acquisition mechanisms" (motivation, attention and rule learning, Lotem and Halpern 2012) to enable them to process the relevant acoustic cues and to respond appropriately. In accordance with this, selection may well have acted on the perceptual system of the green monkey to enable them to both recognise the small but biologically relevant differences existing between chirps, and perhaps also to form the relevant associations faster while experiencing a high level of arousal. An alternative explanation is that signal receivers may respond more strongly to leopard chirps because they are a more urgent call associated with contexts of higher caller arousal, a point supported by the finding that leopard chirps differ from snake chirps in parameters that frequently indicate higher caller arousal (Briefer 2012).

It is important to note that despite their apparent ability to differentiate to some extent between leopard and snake chirps, green monkeys, like vervets (Seyfarth et al. 1980b), do sometimes respond with inappropriate escape behaviour. This could be explained by the unequal costs of inappropriate responses (Godfrey-Smith 1991). For example, the high cost to individuals of not climbing a tree when a leopard is present versus the smaller cost of climbing a tree when a snake is present, could have led to a bias of green monkeys attributing chirps to leopard presence when the signal is ambiguous in terms of its association with either a leopard or snake. That green monkeys in this study sometimes responded to snake chirps by climbing into a tree supports this hypothesis, but the finding that subjects did not always respond to leopard chirps by climbing into a tree does not. It could also be that climbing into a tree is, in some cases, an adaptive response to a snake, and/or that other contextual cues are required for listeners to attribute meaning to their chirps with a high degree of certainty.

In this study, priming with a leopard model increased the chances of a subject responding to both leopard and snake chirps with a leopard typical response (climbing >2m into a tree), but

this effect was small and did not reach significance. It is possible that the effects of context on such responses are subtle and were not picked up with the small sample size of this study. It could also be that the contextual prime (presented up to an hour before the playback of calls) became less relevant over longer time intervals. This could explain differences between this and another study in which context was found to affect Diana monkey responses to Guinea fowl alarm calls, as contextual primes in that study were given just 5 minutes prior to the broadcasting of alarm calls (Zuberbühler 2000c). However, the behaviour of vervet monkeys indicates that they remember the location of a predator for at least 2 hours after seeing it (Cheney and Seyfarth 1992). It is known that vervets frequently respond to playbacks of predator-specific alarm calls by looking towards the speaker and scanning the surrounding environment before responding with escape behaviour (Seyfarth et al. 1980b), and laboratory studies suggest that a subject's surroundings can affect how conditioned stimuli are perceived (Pearce and Bouton 2001). It is therefore possible that contextual cues present at the time of call perception (e.g. the behaviour of group members), were more salient than the recent sighting of a predator and had a larger influence on listeners' initial attribution of meaning and immediate response.

In contrast to the lack of an effect of priming context on immediate responses, we did find that both call type and context had an effect on the length of time that subjects remained in a tree following a playback. Specifically, subjects stayed longer in a tree after hearing leopard chirps for the most part only after having been primed with a leopard model. The lack of an interaction between call type and context is likely due to the fact that GLMMs lack the power to identify interactions when sample sizes are small (R. Mundry, personal communication). It is therefore possible that an individual's prior knowledge was incorporated to refine meaning attribution at a later point in time, leading to the individual staying longer in a tree when both vocal and contextual cues pointed to a leopard being present. Alternatively, signal perception may involve separate meaning attribution and decision making processes, each of which may vary based on additional contextual cues (Fischer 2013). If this is the case, it is possible that staying longer in a tree was the result not of a difference in meaning attribution, but of a difference in a subsequent decision making process.

## Conclusion

Adult green monkeys respond to graded differences in the vocal structure of their chirp calls, on average, with an appropriate anti-predator escape behaviour. The fact that acoustic

cues were insufficient to elicit appropriate responses in all individuals, however, suggests that context likely does play a role in how green monkeys attribute meaning to these calls, but that a receiver's prior knowledge may play a role in delayed rather than immediate attribution of meaning. Studies that systematically test, under natural settings, whether different types of contextual cues are integrated as a part of meaning attribution and/or feed in to a separate decision making process will be particularly useful in furthering understanding of the flexibility of cognitive mechanisms underlying call perception in animals.

## **Acknowledgements**

We thank the *Diréction des Parcs Nationaux* and *Ministère de l'Environnement et de la Protéction de la Nature de la République du Sénégal* for permission to work in Niokolo Koba National Park, and Jacky Bassene, Cheikh Sane, Samba Ciss, Assane Falle, Nele Fröse and Matthias Klapproth for their assistance in data collection. Thanks to Urs Kalbitzer for sharing information and predator models from his previous study of the green monkeys, and to Laura Almeling, Andrea Schell, Tanya Wolf, and Nora Lindstrom for help with video analyses. Roger Mundry kindly provided statistical assistance and Brandon Wheeler gave helpful comments on the manuscript. TP thanks Kate Arnold for discussions of primate inference during their time together in the forests of Nigeria, and Charlotte Price, Stephen Bradshaw, Oumar Ndiaye, Janna Etz and Nele Fröse for their help constructing predator models. The study was supported by funding from the *Deutscher Akademischer Austauschdienst* and *The Leakey Foundation*.

## **Ethical standards**

The experiments described in this paper comply with the current laws of Senegal and Germany, and were conducted with the agreement of the *Diréction des Parcs Nationaux* and the *Ministère de l'Environnement et de la Protéction de la Nature de la République du Sénégal* (Permit numbers: 0383/24/03/2009; 0373/10/3/2012).

## **Conflict of interest**

The authors declare that they have no conflict of interest.

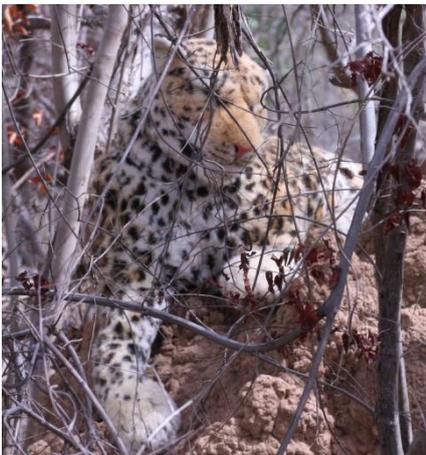
## Supplementary

### S4.1 Predator simulations and modes of presentation

Snake models consisted of one rubber snake painted to resemble a black mamba and two python models. Leopard models consisted of three life-size stuffed toy leopards and one leopard print fur cover. For one experiment looking at behavioural responses to predators, an audio presentation of leopard growls was used. Pictures of a selection of these predator models are shown below.

To simulate snake presence, a snake model was hidden in the undergrowth 11-26m from the subject and drawn across the ground using translucent fishing twine and a custom built pulley system. To simulate leopard presence, a leopard model, or a person draped with leopard cover was hidden 7- >60m from the subject until the time of the experiment when the model or person was revealed. For one experiment an audio speaker was concealed 20m from the subject and leopard growls (recorded by TP from a local captive leopard) were broadcast.

**Leopard model**



**Snake models**



**S4.2 Chirp playback stimuli: Statistical analysis**

**Table S4.2** Transformations applied to the raw data of acoustic parameters prior to LDA

| Parameter                | Transformation  |
|--------------------------|-----------------|
| Duration                 | $\wedge^{-0.2}$ |
| Frequency range          | $\wedge^{0.1}$  |
| Jump                     | Log             |
| Peak frequency deviation | Log             |
| Second quartile_1        | $\wedge^2$      |
| Second quartile_2        | Square root     |
| Second quartile_3        | Log             |
| Second quartile_4        | Log             |
| Third quartile_1         | $\wedge^{-0.6}$ |
| Third quartile_3         | $\wedge^{-0.6}$ |
| Third quartile_4         | $\wedge^2$      |
| Peak frequency_2         | Log             |
| Peak frequency_3         | Log             |
| Peak frequency_4         | Log             |



## **Chapter 5**

### **General Discussion**

Within this general discussion I will summarise empirical results concerning features of vocal flexibility, semantics and pragmatics within African green monkey vocal communication in relation to the underlying mechanistic (proximate) and evolutionary (ultimate) causes. Within this framework I will compare my findings to what is currently known about the proximate mechanisms underlying call production and call perception in primates and other animals, and describe how appraisal theories used to investigate human emotion could be applied to describe and further assess the interaction between cognitive and emotional components of vocal communication in animals. Following this I will discuss the contributions this study makes to an understanding of the ultimate causes and consequences of vocal variation in primates with regards to phylogeny and function, before considering the overall implications of my findings in relation to potential evolutionary stages in the emergence of language. Finally I will provide a short conclusion and outlook for future research that could fill current gaps in our understanding of vocal communication in animals with the aim of achieving a clearer picture of the selective pressures shaping animal communication, and how language and emotional prosody emerged during homonin evolution.

#### **5.1 Proximate Mechanisms**

##### **5.1.1 Call production**

Language is based upon a system of flexible vocal production and semanticity, which is expressed in the geographic diversity of human speech and arbitrary structure of words. Objectives of this thesis concerning proximate mechanisms of call production were thus to first to assess geographic variation in African green monkey barks in order to describe and quantify vocal flexibility within this genus; then to evaluate the relationship between the structure of African green monkey calls and external eliciting stimuli with the view to offering insights into the emergence of linguistic semanticity.

### Flexibility

In **Chapter 2** I show that, in contrast to the learnt structural variability observed in human speech and bird song (Doupe and Kuhl 1999), geographic variation in the spectral structure of adult male African green monkey barks was subtle, and could be attributed to phylogenetic differences. From this finding I drew the conclusion that call production within this genus is largely an innate process, although further study of intra-species population differences in fine spectral structure would strengthen this claim. Generally vocal production in primates is strongly constrained (reviewed in Hammerschmidt and Fischer 2008), and a genetic basis for intra- and inter-species vocal variation has been proposed across a wide range of primate taxa (*Leontopithecus* spp., Snowdon et al. 1986; *Lepilemur* spp., Méndez-Cárdenas et al. 2008; *Presbytis thomasi*, Wich et al. 2008; *Nomascus* spp., Thinh et al. 2011; *Presbytis* spp., Meyer et al. 2012). It is therefore not surprising that African green monkey barks comply with this general pattern. This constraint on call production can most likely be attributed to a lack of direct connections between motor cortical areas and motor neurons controlling laryngeal movements (Jürgens 1976a), the connection which underlies the production of novel sounds in humans (Kuypers 1958). This supports Darwin's (1872) premise that differences in vocal production between man and animal are for the most part related to differences in neural connections rather than vocal anatomy.

That bark duration was the most distinctive feature to differ between African green monkey species also fits well with current understanding of primate vocal neurobiology, as during the production of innate calls, the PAG plays a role in gating call initiation and influencing intensity but not patterning (Jürgens 1992), with the implication that primates have more control over temporal than spectral characteristics, so that temporal features are likely to be less evolutionarily constrained. It is thus interesting that acoustically similar vervet calls produced in predator and non-predator contexts could often be discriminated by differences in temporal parameters (**Chapter 3**), as this suggests that selection may act on temporal features when adaptive responses require receivers to differentiate between spectrally graded calls. This also fits with the structural differences identified between graded green monkey chirps (**Chapter 4**), whereby the duration of call elements was the most influential parameter when distinguishing between leopard and snake chirps.

### Semantics

In **Chapter 3**, reanalysis of vervet alarm calls supports the findings of earlier studies (Struhsaker 1967b; Seyfarth et al. 1980b) that these calls are distinguishable with regards to the predator class eliciting calling. Important new insights are also offered, in that spectral

differences between call elements produced in response to different predator types, and between call elements produced in predator and non-predator contexts are often of a graded nature. As outlined in the general introduction, graded differences in vocal structure are frequently attributed to variation in the caller's arousal level, brought about by changes in the autonomic and somatic nervous system and subsequent modification to vocal production anatomy (Briefer 2012). In accordance with broad patterns identified in other species, differences in peak frequency observed between some vervet alarm and acoustically similar non-alarm calls could be the result of the caller experiencing higher arousal in predator than non-predator contexts. Similarly, adult female and juvenile green monkeys were found to produce longer and higher frequency chirps in response to leopards than to snakes (**Chapter 4**), a finding indicative of higher caller arousal in the presence of leopards. Temporal differences in the number of calls produced and rate of calling were harder to interpret; a more objective method of defining temporally clumped calls would have been advantageous, but this is often a problem with field recordings when it is difficult to pick out all incidences of an individual's calls. Incorporating findings from neurological studies on PAG firing patterns could help to objectively define call categories from the caller's perspective. In addition, how temporal parameters vary in relation to caller arousal appears to differ between species, for example in some species call duration has been found to be longer in contexts of supposed higher arousal (*M. sylvanus*, Fischer et al. 1995; *Saimiri sciureus*, Fichtel et al. 2001; *Canis lupus*, Yin and McCowan 2004; *P. troglodytes*, Slocombe and Zuberbühler 2007; *L. africana*, Soltis et al. 2011), whilst in other species the reverse pattern has been observed (*Suidae sus*, Schrader and Todt 1998; *P. ursinus*, Fischer et al. 2001a; *S. suricatta*, Manser 2001). Such trends may also differ between different call types.

While the graded vocal differences identified in this study are strongly suggestive of arousal-mediated effects in call production, ultimately more direct measures of caller arousal, such as stress hormones or heart rate will be required to directly assess vocal correlates of anxiety in African green monkeys. Indeed this is also more broadly the case across species, as vocal correlates of stress are frequently assessed based on inferences of how stressful a situation is (Jovanovic and Gouzoules 2001; Soltis et al. 2011; Collins et al. 2011), and studies using direct measures often look at call occurrence rather than differences in call parameters (Boinski et al. 1999; Byrne and Suomi 1999; Norcross and Newman 1999; but see Schrader and Todt 1998; Marchant et al. 2001; Byrne and Suomi 1999; Norcross and Newman 1999).

The results of **Chapter 3** also support earlier findings that vervet monkeys produce different call types to different predators, nevertheless, they further suggest that when male alarm calls are considered separately and when non-alarm calls are incorporated within analyses, vervet

alarm calls demonstrate lower context specificity than would be expected for an arbitrary relationship between a call and the eliciting external stimulus (**Chapter 3**). One explanation, in line with the early hypotheses of Darwin (1872) and Morton (1982), is that leopards, snakes and raptors induce different motivational states that in turn elicit different vocal types and associated somatic responses, whereas callers experience similar motivational states in some predator and non-predator contexts. For example males may produce barks to leopards and during aggression as a result of experiencing an aggressive drive to display. Support for this hypothesis comes from neurological studies showing that during vocal production in squirrel monkeys, and likely all other primates, the PAG predominantly receives input from various limbic pathways and triggers the initiation of different species-specific call types on the basis of this motivation-related information (Dujardin and Jürgens 2005). Additional support comes from studies showing that electrical and chemical brain stimulation that elicits positive or aversive emotional states in rats, chickens, cats, and squirrel monkeys also elicits the production of distinct vocalisations (Jürgens 1979; Normansell and Panksepp 2011; Brudzynski 2013); and that administration of benzodiazepine compounds (drugs with anti-anxiety effects) reduces defensive and attack behaviours and defensive vocalisations in rats, mice, cats, and primates (reviewed in Blanchard et al. 1998). On the other hand, a recent study carried out by Hage and colleagues (2013) report flexible production of two different call types to arbitrary visual prompts in a trained rhesus macaque (*M. mulatta*), and the authors interpret this as showing that primates can decouple their innate vocalisations from the emotional state experienced. An important limitation of this study is that it focused on call types produced naturally in response to food, and the subject was rewarded with food after each test trial; it is therefore difficult to exclude the possibility that an emotion-based anticipation of food award played a role in call production. Nonetheless, the ability to switch between two different call types is suggestive of a higher level of control than previously supposed, and warrants further study to understand the underlying neural mechanisms.

In the above paragraphs I describe how African green monkey vocalisations could be influenced by both caller arousal and motivational state, highlighting the point that one-dimensional differences in arousal are unlikely to capture the complexities of a caller's emotional state (Jürgens 1979). The "biphasic theory of emotion" that integrates a two dimensional approach (including arousal and valence; Lang et al. 1998), or a three dimensional framework (including arousal, valence and potency; Goudbeek and Scherer 2010) may better predict the relation between behaviour (including vocalisations) and emotional state. Mendl and colleagues (2010) describe the advantages of bringing discrete and dimensional approaches together for future study of animal emotion. Appraisal theories of human emotion take a similar perspective,

and in section 5.1.3 I introduce appraisal theories as a useful framework to better understand the complexity of emotional events in animals, and to investigate the important question of how cognitive and emotional components may interact to influence vocal behaviour (Fischer et al. 2001a).

### **Summary**

Patterns of call production in African green monkeys are most parsimonious with the conclusion that vocal structure is innate, with more flexibility in temporal than spectral characteristics. Call structure is likely influenced by components of the caller's emotional state. This interpretation supports the growing consensus that animal vocalisations are more similar to human innate and emotionally-grounded non-verbal sounds, such as cries and laughs, than to speech (Owren et al. 2011; Fitch and Zuberbühler 2013). Importantly, this does not imply that vocal production is a reflexive or involuntary act, or that higher cognitive processes do not play a role in regulating or triggering distinct emotional states, a concept central to appraisal theories of emotion.

### **5.1.2 Call perception**

When humans hear a word or phrase they tend to infer meaning from the arbitrary conventional relationship that exists between words and external referents (semantics-related meaning), and from the context in which the utterance was made (pragmatics-related meaning). Objectives of this thesis concerning proximate mechanisms of call perception were thus to investigate whether green monkeys attribute meaning to their alarm calls, and whether contextual cues in the form of previous experience influence this process.

### **Semantics**

In **Chapter 4** I show that green monkeys were more likely to respond to leopard-chirps than to snake-chirps by climbing more than 2m up into a tree, a typical leopard avoidance behaviour as established previously with predator-model presentation experiments. This finding suggests that green monkeys may be capable of categorical perception, although further study using a habituation-dishabituation paradigm (described in Fischer 2013) is required to assess how intermediate vocal structures are perceived. Categorical perception was once thought to be unique to humans (Liberman and Pisoni 1977), but it is now recognised as being widespread within the animal kingdom (insects: Wytenbach et al. 1996; amphibians: Baugh et al. 2008; birds: Nelson and Marler 1989; mammals: Kuhl and Miller 1975; Morse and Snowdon 1975). An interesting finding is that while Barbary macaques respond to conspecific's graded calls

categorically (Fischer 1998), chacma baboons ignored intermediate call variants and may pay more attention to contextual than acoustic cues (Fischer et al. 2001b). How different sources of information are integrated is discussed under the heading of pragmatics just below.

That green monkeys respond appropriately to leopard-chirps could be interpreted as green monkeys having learnt to associate leopard-chirps with leopard presence, or equally it could be explained by affect-conditioning (Owren and Rendall 1997) if green monkeys experience a more fearful state when hearing leopard-chirps than snake-chirps and if tree-climbing behaviour is related to the degree of fear experienced. In some species, studies have shown that subjects look up in response to aerial alarm calls (*Saguinus fuscicollis* and *S. mystax*, Kirchof and Hammerschmidt 2006; *Callicebus nigrifrons*, Cäsar et al. 2012). This is more supportive of a representational interpretation, and this is in keeping with the finding that aerial alarm calls often show the highest degree of predator-specificity (Zuberbühler et al. 1997; Fichtel and Kappeler 2002; Wheeler 2010). For calls with lower predator-specificity as is often the case with terrestrial/general alarm calls, learnt associations between calls and motivational states, and/or the incorporation of contextual information (Wheeler and Fischer 2012) may well be a more adaptive solution. Experimentally manipulating the salience of different predator types has the potential to offer more conclusive support for representation-based associative learning (Evans 1997); for example if subjects associate a leopard alarm call with a leopard rather than with a state of fear, habituating subjects to leopards should subsequently reduce responses to the leopard-associated call. A study like this was carried out on Diana monkeys (Zuberbühler et al. 1999a) and supported a representational conclusion; longer temporal separation of habituation and stimulus presentation would, however, be beneficial to more strongly negate an affective explanation.

Vervet barks also provide cues to caller identity at the individual and species level, and responses to playback experiments indicate that adult males distinguish between known and unknown conspecific calls, and conspecific and heterospecific calls (**Chapter 4**). These are differentiations that are highly relevant in a territorial species living in multi-male groups, in keeping with the prediction that individuals' social attention corresponds to the specifics of the species' social organisation (Maciej et al. 2013). Earlier studies suggest that vervets are also capable of more complex vocal recognition, recognising third party relationships (Cheney and Seyfarth 1980) and which neighbouring group a caller belongs to (Cheney and Seyfarth 1982a).

### **Pragmatics**

Looking at the incorporation of contextual cues, **Chapter 4** also shows that priming condition did not have a strong effect on green monkeys' immediate responses to chirp

playbacks, but did influence the amount of time that the individual stayed in a tree before descending to the ground again. Bees (*Apis mellifera*) and ants (*Lasius niger*) also make use of private contextual information (that is, previous experience) when making foraging decisions (Grüter and Ratnieks 2011; Grüter et al. 2011), but incorporate such cues in immediate rather than subsequent responses. That foraging decisions are likely less urgent than responses to alarm calls may have an influence on how contextual cues are incorporated within decision-making events, but private contextual cues also influence the immediate responses of Diana monkeys (Zuberbühler 2000c) and putty nosed monkeys (*Cercopithecus nictitans*, Arnold and Zuberbühler 2013) within predator-contexts. In these last two experiments priming conditions were given just minutes or seconds prior to the test stimulus, however, which may have meant the primed information was more easily retrieved than in my study when priming occurred up to an hour before. Alternatively, the findings of these studies could be viewed within the framework of affect-based cognitive bias, whereby priming induced a change in subjects' emotional state which subsequently influenced the processing of test stimuli; I discuss affect-based cognitive biases in more detail below, but an important point to make here is that, within this scenario, duration between prime and test stimuli would likely be highly relevant. These descriptions highlight that whilst the incorporation of contextual cues is likely widespread, little is known about species and individual differences in the type of contextual cues that are incorporated and the level of automaticity at which they are processed.

### **Summary**

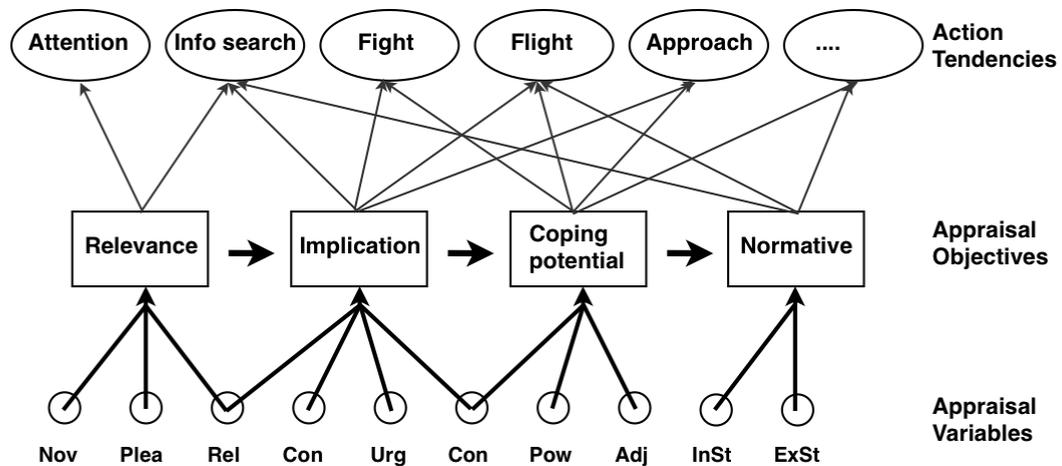
Learnt associations likely underlie green monkeys' and vervet monkeys' responses to conspecific chirps and barks, which may be based upon a representation of the stimulus or affect-conditioning. Context also influences later response behaviour. As with call production, there are still many unanswered questions concerning the cognitive processes involved in information processing, and how this interacts with more basal emotional states. A theoretical starting point for future investigations is that the meaning inferred by a signal receiver and their subsequent responses could be viewed as two separate processes, and that context could play a different role in both (Fischer 2013). I propose that, in addition to providing a framework for understanding call production in animals, appraisal theories could help to disentangle the processes involved in the perception of, and responses to, these calls.

### 5.1.3 Appraisal theories of emotion

Appraisal theories, first proposed by Arnold (1960) and Lazarus (1966), consider human emotion as adaptive changes in an individual's motivational state, peripheral physiology, expressive behaviour, and feeling, which are typically brought about by various assessments of the environment that are relevant to the individual's well-being (Lazarus 1991; Roseman and Smith 2001; Scherer 2001; Moors et al. 2013). Scherer's component process model (CPM), one class of appraisal theories, describes the appraisal event as first eliciting a motivational effect, which in turn brings about changes in the individual's autonomic and somatic nervous system (including respiration and heart rate) and expressive behaviour (including vocal output) which both affect each other and feed back into re-appraisal processes (Scherer 2009a). This fits well with the current picture painted by neurological studies of vocal production in animals, which describes motivationally-triggered genetically pre-programmed vocal patterns (Dujardin and Jürgens 2005) that are subsequently modified by changes in caller arousal in relation to changes in the autonomic and somatic nervous systems (Scherer 2003). In the following paragraphs I describe how CPM could provide a framework for investigating the cognitive appraisal process in animals, including the assessment of vocal cues, and the interplay between this and behavioural and vocal responses.

The first point of interest of the CPM model is that it specifies a number of appraisal variables that feed into overarching appraisal objectives, which in turn influence the individual's motivational state and subsequent behaviours (Moors et al. 2013). Appraisal objectives and constituent variables constitute "stimulus evaluation checks" (SECs; Scherer 2001), and Figure 5.1 provides a schematic diagram of this process and the names of proposed SECs. As illustrated, the CPM model proposes that stimulus evaluation is based on four appraisal objectives that pertain to stimulus relevance (whether the stimulus requires attention, action or further information processing), stimulus implications (pertinent characteristics of the stimulus and consequences for the individual), coping potential (the individual's ability to cope with the expected consequences), and normative significance (assessment of event in relation to social norms and/or moral rules). Taking relevance as an example, this can be broken down into three appraisal variables: novelty (the likelihood of occurrence), intrinsic pleasantness (how pleasant the stimulus is irrespective of the individual's motivational state), and goal relevance (is the stimulus relevant with regards to current motivational state). Appraisal variables such as agency and internal standards are unlikely to be relevant for animal studies. Nevertheless playback studies indicate that some forms of "social norms", (i.e. expectations based on species-typical features of the social structure and mating behaviour) are integrated within primate decision making (Bergman et al. 2003; Crockford et al. 2007). This list of variables therefore provides a

good starting point to test predictions of species and ontogenetic differences in which SECs are relevant when producing or responding to vocal signals.



**Figure 5.1** Schematic layout of stimulation evaluation checks (SECs) and resultant action tendencies; modified from Scherer (2009a). Nov=novelty, Plea=intrinsic pleasantness, Rel=goal relevance, Con=conduciveness to goals, Urg=urgency to react, Con=controllability of event, Pow=personal power to exert control, Adj=adjustment to consequences, InSt=compatibility with internal standards, ExSt=compatibility with external standards

A second useful application of appraisal theories to animal emotion could be in formulating and testing hypotheses related to the underlying cognitive mechanisms and degree of automaticity involved during SECs. Leventhal and Scherer (1987) describe three different levels at which appraisal processes could be carried out. As laid out by Scherer (2009a), the simplest mechanism in terms of cognitive processing is at the sensory-motor level, which consists of species-specific genetically-determined templates for pattern matching. The intermediate mechanism is at the schematic level and is based on social learning processes that are relatively automatic and unconscious. The most cognitively complex mechanism is at the conceptual level, which incorporates processing within cortical association areas and requires propositional knowledge. Importantly the level of processing likely depends on the novelty of the stimulus (Ellsworth 2013) and which SEC is being carried out (Scherer 2013), which means that a single event can be processed at different levels depending on the experience of the individual and which SEC is being processed. Hypothesised order of processing is illustrated within Figure 5.1 with arrows moving from Relevance through to Normative appraisal objectives (Scherer 2009a). Support for such temporal processing differences in humans comes from neurological studies which have identified different brain areas and different processing times for different SECs; in

particular novelty detection appears to be a rapid and automatic process taking place in the amygdala and hippocampus, whilst the assessment of goal congruence occurs later and perhaps less automatically and takes place in the cingulate and prefrontal cortex (reviewed in Scherer 2009b; Brosch and Sander 2013).

It may be that more immediate SECs are the equivalent of Fischer's and Wheeler's (Wheeler and Fischer 2012; Fischer 2013) description of attribution of meaning, whilst SECs processed at higher levels relate to their description of decision-making processes. Nevertheless, appraisal theories are based upon the surmise that appraisal is a constant recursive process within which the outcomes of each cycle are fed back to the appraisal component for re-appraisal (Ellsworth and Scherer 2003), suggesting that meaning attribution is best viewed as a continual process. This highlights the important point that appraisal theories consider the relation between cognitive processing and emotional states to be bidirectional. The influence of emotional state on cognitive processing can be described as "affect-induced cognitive bias", (Mendl et al. 2010), and has been demonstrated in some species in captivity (reviewed in Mendl et al. 2009). An understanding of whether cognitive bias influences vocal communication in the wild (as suggested in discussion of pragmatics in section 5.1.2), is a promising field of future research (Semple and Higham 2013). To better understand how meaning-attribution and decision-making processes can be differentiated and related to specific SECs, hypotheses can be formulated to test whether similar patterns of brain activation (areas and timing) are found in animals when vocal and vocal-eliciting stimuli, and individual's motivational state are manipulated (Fischer 2013); providing a tool to assess parallels between the expression and perception of emotional cues in the human and animal voice (Fischer 2011). In addition, appraisal theories of emotion assume a degree of functionality, based on the surmise that emotions are adaptive and function to bring about responses to the environment that are relevant to the individual's well being (Ellsworth 2013). Thus viewing vocalisations as part of this emotional state provides a clear link between proximate theories of emotional mechanisms underlying call production, and ultimate theories of vocalisations being shaped by functional demands.

## **5.2 Ultimate Mechanisms**

### **5.2.1 Evolution of alarm calls**

Current theory concerning the evolution of distinct alarm calls is that these calls have arisen in response to the need for qualitatively different reactions to predators with different hunting strategies (Macedonia and Evans 1993). The vervet alarm calling system is often proposed as

strong support of this hypothesis, because the spectral and temporal characteristics of chutters, chirps, barks and rraups make each call type well-suited to bring about an appropriate response with regards to the hunting strategy of the predator class which typically elicits them. Nevertheless an alternative explanation is also possible. In Struhsaker's (1967b) early descriptions, chirps and barks are described as being produced to more than one predator class, a finding supported by my re-analysis of vervet barks (**Chapter 3**) and analysis of green monkey chirps (**Chapter 4**), and, at the element level, vervet chutters, rraups and barks produced in predator contexts grade into calls produced during social and aggressive interactions (**Chapter 3**). Considering these findings together, it is also plausible, therefore, that vervet alarm calls did not evolve in response to different predation pressures, but rather were co-opted from social and/or aggressive contexts. From the current data it is not possible to differentiate between these hypotheses, rather phylogenetic comparative studies are needed to determine whether each vervet alarm call arose in response to predation pressures or was co-opted from other non-predator contexts. With regards to male barks, interesting insights are offered by a comparative study of long calls in primates that found that the presence of long calls is the ancestral state, and that they have frequently been lost in terrestrial, savannah-living males perhaps due to an increased need for within-group than between-group male competition (Wich and Nunn 2002). On the basis of this finding, it is possible that the bark of male African green monkeys, and also the wahoo of male chacma baboons are the descendants of the ancestral male long call, adapted first for within group male competition and then co-opted for defence against terrestrial predators.

Irrespective of the evolutionary trajectory of vervet alarm calls, results from **Chapter 3** suggest that selection has acted on females more than males to produce structurally distinct calls in response to different categories of predators. This differs from species such as the white face capuchin (*Cebus capucinus*, Fichtel et al. 2005) in which males and females produce the same predator-specific alarm calls, species such as the putty nosed monkey (Arnold and Zuberbühler 2006a) in which males have a larger repertoire of predator-elicited calls than the female, and species such as the chacma baboon (Fischer et al. 2001a; Fischer et al. 2002) in which males and females both produce a single sex-specific alarm call, with graded subtypes related to different predator and non-predator contexts. Vervet monkeys live in philopatric multi-male multi-female groups with a distinct breeding season (Struhsaker 1967c; Fedigan and Fedigan 1988) so that females have more relatives within the group, and the alpha-male is unable to monopolise matings (Cheney and Seyfarth 1992). Taking these social factors into account, sex differences in vervet alarm calls could be explained if producing alarm calls increases female fitness by increasing both individual survival (through mobbing: Curio 1978; or

predator deterrence: Hasson 1991) and the survival of close kin (by warning kin or defending offspring: Hamilton 1963; Hamilton 1964) through appropriate responses to different predator classes; but increases male fitness primarily by increasing reproductive success not by offspring survival but via mate attraction and/or male-male competition. Playback experiments carried out on adult male vervets (**Chapter 2**) support the claim that barks have a dual function as both alarm and display call. Looking more broadly at the influence of social factors on vocal production in animals, social complexity has been proposed as a driver of communicative complexity (Freeberg et al. 2012), and is related to vocal repertoire size across a wide range of primates (McComb and Semple 2005), and to the number of alarm calls within the vocal repertoire of marmots (Blumstein 2007). The diversity of alarm calling systems found only within the subfamily *Cercopithecinae*, however, emphasises the need for broad comparative analyses to tease apart the influence of social system, phylogeny, predation pressure and environment on the evolution of context-specific alarm calls in primates (Fischer 2013).

Within this section I have so far focused on factors affecting the evolution of different types of alarm call; I will now briefly discuss the evolutionary factors underlying variation within the vervet alarm bark. Differences in the spectral structure of male African green monkey barks could be attributed to genetic distance between populations, and the lower frequency bark of the green monkey is likely the result of the western-species' larger body size (**Chapter 2**). Interestingly, patterns of skull size within the genus map to clinal variation in rainfall, and it has been hypothesised that larger body size is influenced by differences in habitat productivity (Cardini et al. 2007). Under this prediction, variation in the spectral characteristics of barks would follow rather than drive genetic divergence. Another possibility is that body size and related vocal characteristics are "magic traits", in that they are simultaneously under divergent selection and contribute to non-random mating (Servedio et al. 2011). This process can lead to rapid diversification (Wilkins et al. 2013) which is a feature of early lineage divergence between *Chlorocebus* taxa (Haus et al. 2013). Further investigation would be needed to test this prediction, however, which could be done by assessing whether low frequency calls are related to male reproductive success to a higher degree in green monkeys than in vervets.

The duration of Ex1 bark elements was the most influential parameter distinguishing between green monkey and vervet barks, with green monkeys producing longer call elements than vervets (**Chapter 4**). Acoustic properties such as call duration are in some cases better explained by energetic than physiological constraints, and in some other primate species, higher ranked individuals (with rank taken as a proxy of caller fitness) have been found to produce longer calls (*P. ursinus*, Fischer et al. 2004; *Macaca nigra*, Neumann et al. 2010). Observations and predator-model presentations with East African vervets suggest that higher ranking males

are more likely to initiate calling and to call for longer (Cheney and Seyfarth 1981). It could be that in green monkey populations longer call durations are a sexually selected trait, with the hypothesis that higher ranking green monkeys would produce longer duration call elements. Unfortunately I do not have the data to test this, but the suggestion is made more likely in light of the finding that the bark appears to have a dual function as both an alarm and a display call. Importantly, the dual function of barks may well account for the degree of geographic variability observed, which was greater than calls produced exclusively in alarm contexts, but less than calls with a primarily display function. This then supports the prediction that the strength of stabilising constraints is related to call function (Struhsaker 1970), and that sexual selection plays an influential role in selecting for variation within the structure of innate calls, as well as being a potential driver for the evolution of vocal learning through mating preferences for varied vocalisations (Nottebohm 1972; Jarvis 2006).

### **5.2.2 Evolution of language**

Current theories of how modern language emerged describe a process of gradual evolution, with at least one intervening stage or "protolanguage"; nevertheless, the characteristics of proposed protolanguages differs between theories (Fitch 2010). In the following few paragraphs I will take a broad perspective of the findings presented in this thesis, and of the wider implications discussed above, to give my opinion on the insights they offer into potential intervening stages during the evolution of language.

One class of theories propose a "lexical" protolanguage, whereby at one point in our evolutionary past our hominin ancestors possessed a large, learned lexicon of meaningful words without complex syntax (Bickerton 1990; Jackendoff 1999). Within this thesis I do not consider the evolution of syntax, thus questions of interest here are related to the earlier stage, how did learned and meaningful words emerge? With regards to this question, the functionally referential calls of primates have received much attention as they have been proposed as a potential precursor to symbolic meaning (Zuberbühler 2003; Fedurek and Slocombe 2011). Proponents and opponents of this hypothesis note important caveats, however, which the findings in this thesis support. Namely, that functional reference in primates depends on an indexical relation between a small number of innate and functionally constrained (i.e. non-arbitrary) context-specific calls, and a temporally and/or spatially related referent (Wheeler and Fischer 2012), and it is unclear how these constraints could have been circumvented to evolve into the large repertoire of arbitrary calls upon which the proposed lexical protolanguage is based; indeed it has been suggested that indexical communication systems (such as functionally

referential calls) may even block the emergence of symbolic communication (Deacon 1997). An additional important point is that associative learning is found in all major taxa and across a wide range of contexts (Heyes 2012). Viewed in this light, functional reference is unlikely to have emerged as a unique adaptation for referential communication, and learnt associations between referents and functionally referential signals are logically less complex than those between referents and less context-specific calls (Wheeler and Fischer 2012). Nevertheless the capability to form arbitrary concepts may well have been an exaptation or "spandrel" (Gould and Lewontin 1979) that was put to use by the developing language faculty, as "no biologist would expect every aspect of a trait as complex as language to represent an adaptation" (Fitch 2010, p. 66).

Musical protolanguage provides an interesting alternative hypothesis, originating back to Darwin (1872) who described generative vocal production as emerging prior to it being put to meaningful use. Updated and extended by multi-disciplinary scholars, such as Jespersen (1922), Wray (1998) and Kirby (1999), the following description is based upon Fitch's (2010) four-stage synopsis of how language emerged from the vocal communication system of our last common ancestor with chimpanzees. During a first stage, the acquisition of complex vocal learning brought about an initial song-like phase, for which sexual selection was likely an important driver. An important question is how direct cortical control over vocal production was evolved, and one hypothesis is that expansion of the human neocortex may have played a key role in this development, by allowing vocal learning pathways to compete more successfully with the otherwise dominant neural pathways underlying innate call production (Deacon 1997). In a second stage arbitrary, holistic meaning arose from an ability to form associations between whole phonological signals (which at this stage were still emotionally grounded manipulative signals) and whole semantic events. Interestingly, this concept of holistic meaning from whole emotionally grounded phonological signals provides a potential parallel to meaningful call sequences that have been identified in primates (Zuberbühler 2002; Clarke et al. 2006; Arnold and Zuberbühler 2006b), suggesting that some features of syntax such as hierarchy and concatenation were present in the communication systems of our primate ancestors. During a third stage, analytic meaning was acquired through linked wholes being broken down into smaller meaningful lexical pieces, and in a fourth stage, modern language appeared, perhaps together with the emergence of theory of mind (Malle 2002). While there are still many unanswered questions, and many untested predictions, this theory fits well with the finding that across many species, sexual selection is a primary driver of vocal variation, it presents a scenario in which pre-existing capabilities for associative learning and concept formation (but not Gricean intentions) are adopted by the language faculty, and it provides a plausible explanation for the

origin of arbitrary signals from an innate, emotionally grounded and functionally constrained system.

### **5.3 Conclusion and Outlook**

In accordance with findings across the primate order, vocal production in the African green monkey appears to be largely innate and shaped by functional constraints. Future research should aim to shed light on the selective pressures shaping vocal communication in primates, and to better understand both the influence of sexual selection on vocal variability within primate species and the constraints imposed by vocally-mediated species recognition. Building on findings from this thesis, interesting questions are whether spectral and/or temporal features of the male African green monkey bark correlate with male fitness, and if trends are the same across species; also whether acoustic variation across the African green monkey genus, including hybrid zones, concords with genetic distance, and how this affects vocally mediated species recognition. Comparative studies carried out to identify consistent differences in the degree of geographic variation found for display, social and alarm calls across a wide range of primate and other species would further understanding of the relation between call function and vocal production flexibility.

Variation in the structure and type of call produced by African green monkeys may be related to changes in the caller's emotional state, including differences in motivation state and arousal level. Future research investigating the relation between cognitive and emotional processes underlying call production in animals has the potential to increase our understanding of how animal vocalisations compare to verbal and non-verbal vocal production in humans. Uncovering vocal correlates of emotion in animals also has implications for improving the welfare of animals in captivity. To follow up on my studies with African green monkeys, important next steps would be to investigate whether direct measures of caller arousal, such as salivary glucocorticoid levels, correlate with the production of different call types, and to assess how vocal production varies with external stimuli above the level of the call element, for example at the level of multi-element calls and throughout whole calling bouts. Furthermore, broad comparative studies carried out to manipulate variables relevant to different stimulus evaluation checks, and to assess how this affects call production would increase knowledge of inter-specific differences in appraisal processes potentially underlying call production.

Lastly, African green monkeys demonstrate higher flexibility on the side of call perception, differentiating between calls with a graded structure, forming learnt associations between acoustic cues and features of predator type and caller identity, and integrating contextual cues into behavioural response decisions. Testing whether reducing the salience of leopards as a

stimulus alters how captive vervets respond to subsequent presentation of a leopard-specific alarm call, investigating the effect of affect-based cognitive bias on responses to intermediate exemplars of graded call types, and assessing species and context differences in the incorporation of contextual cues when responding to vocal signals would all help to further elucidate the cognitive processes underlying call perception in African green monkeys and other non-human species.

---

## References

- Aitken PG, Wilson WA Jr (1979) Discriminative vocal conditioning in rhesus monkeys: evidence for volitional control? *Brain and Language* 8:227–240.
- Akçay E, Wood WE, Searcy WA, et al. (2009) Good neighbour, bad neighbour: song sparrows retaliate against aggressive rivals. *Animal Behaviour* 78:97–102.
- Albon SD, Clutton-Brock TH (1979) The roaring of red deer and the evolution of honest advertisement. *Behaviour* 69:145–170. doi: 10.1163/156853979X00449
- Arnold K, Pohlner Y, Zuberbühler K (2008) A forest monkey's alarm call series to predator models. *Behavioral Ecology and Sociobiology* 62:549–559.
- Arnold K, Zuberbühler K (2006a) The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. *Animal Behaviour* 72:643–653.
- Arnold K, Zuberbühler K (2013) Female putty-nosed monkeys use experimentally altered contextual information to disambiguate the cause of male alarm calls. *PLoS ONE* 8:e65660. doi: 10.1371/journal.pone.0065660
- Arnold K, Zuberbühler K (2006b) Semantic combinations in primate calls. Putty-nosed monkeys rely on two basic calling sounds to construct a message of utmost urgency. *Nature* 441:303.
- Arnold MB (1960) *Emotion and personality. Vol. I. Psychological aspects.* Columbia University Press, Oxford, England
- Baayen RH (2011) *LanguageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics."*
- Bates D, Maechler M, Bolker B (2011) *lme4: Linear mixed-effects models using S4 classes.*
- Baugh AT, Akre KL, Ryan MJ (2008) Categorical perception of a natural, multivariate signal: Mating call recognition in túngara frogs. *Proceedings of the National Academy of Sciences* 105:8985.
- Bergman TJ, Beehner JC, Cheney DL, Seyfarth RM (2003) Hierarchical classification by rank and kinship in baboons. *Science* 302:1234.
- Bickerton D (1990) *Language and Species.* University of Chicago Press
- Di Bitetti MS (2005) Food-associated calls and audience effects in tufted capuchin monkeys, *Cebus apella nigrinus*. *Animal Behaviour* 69:911–919.
- Blanchard DC, Griebel G, Rodgers RJ, Blanchard RJ (1998) Benzodiazepine and serotonergic modulation of antipredator and conspecific defense. *Neuroscience and Biobehavioral Reviews* 22:597–612.
- Blumstein DT (1995) Golden-marmot alarm calls. I. The production of situationally specific vocalizations. *Ethology* 100:113–125.

- Blumstein DT (2007) The evolution, function, and meaning of marmot alarm communication. *Advances in the Study of Behavior* 37:371–402.
- Blumstein DT, Armitage KB (1997) Alarm calling in yellow-bellied marmots: I. The meaning of situationally variable alarm calls. *Animal Behaviour* 53:143–171.
- Boinski S, Gross TS, Davis JK (1999) Terrestrial predator alarm vocalizations are a valid monitor of stress in captive brown capuchins (*Cebus apella*). *Zoo Biology* 18:295–312.
- Boughman JW, Moss CF (2002) Social sounds: Vocal learning and development of mammal and bird calls. *Springer Handbook of Auditory Research* 16:138–224.
- Bouton ME, Westbrook RF, Corcoran KA, Maren S (2006) Contextual and temporal modulation of extinction: behavioral and biological mechanisms. *Biological Psychiatry* 60:352–360. doi: 10.1016/j.biopsych.2005.12.015
- Bradbury JW, Cortopassi KA, Clemmons JR (2001) Geographical variation in the contact calls of orange fronted parakeets. *The Auk* 118:958. doi: 10.1642/0004-8038(2001)118[0958:GVITCC]2.0.CO;2
- Bradbury JW, Vehrencamp SL (1998) Principles of animal communication. Sinauer Associates
- Briefer EF (2012) Vocal expression of emotions in mammals: mechanisms of production and evidence. *Journal of Zoology* 1–20. doi: 10.1111/j.1469-7998.2012.00920.x
- Brockelman WY, Schilling D (1984) Inheritance of stereotyped gibbon calls. *Nature* 312:634–636. doi: 10.1038/312634a0
- Brosch T, Sander D (2013) Comment: The appraising brain: Towards a neuro-cognitive model of appraisal processes in emotion. *Emotion Review* 5:163–168. doi: 10.1177/1754073912468298
- Brudzynski SM (2013) Vocalizations as indicators of emotional states in rats and cats. In: Altenmüller E, Schmidt S, Zimmermann E (eds) *Evolution of emotional communication. From sounds in nonhuman mammals to speech and music in man*. Oxford University Press, Oxford, pp 75–92
- Brumm H, Voss K, Kollmer I, Todt D (2004) Acoustic communication in noise: regulation of call characteristics in a New World monkey. *Journal of Experimental Biology* 207:443.
- Byrne G, Suomi SJ (1999) Social separation in infant *Cebus apella*: patterns of behavioral and cortisol response. *International Journal of Developmental Neuroscience* 17:265–274.
- Cardini A, Jansson AU, Elton S (2007) A geometric morphometric approach to the study of ecogeographical and clinal variation in vervet monkeys. *Journal of Biogeography* 34:1663–1678.
- Cäsar C, Byrne RW, Hoppitt W, et al. (2012) Evidence for semantic communication in titi monkey alarm calls. *Animal Behaviour* 84:405–411. doi: 10.1016/j.anbehav.2012.05.010
- Cäsar C, Zuberbühler K (2012) Referential alarm calling behaviour in New World primates. *Current Zoology* 58:680–697.

- Chappuis C (1971) Un exemple de l'influence du milieu sur les émissions vocales des oiseaux: l'évolution des chants en forêt équatoriale. *Terre Vie* 118:183–202.
- Cheney D, Seyfarth R (1990) Attending to behaviour versus attending to knowledge: examining monkeys' attribution of mental states. *Animal Behaviour* 40:742–753. doi: 10.1016/S0003-3472(05)80703-1
- Cheney DL, Seyfarth RM (1985a) Vervet monkey alarm calls: Manipulation through shared information? *Behaviour* 150–166.
- Cheney DL, Seyfarth RM (1981) Selective forces affecting the predator alarm calls of vervet monkeys. *Behaviour* 76:25–61.
- Cheney DL, Seyfarth RM (1982a) Recognition of individuals within and between groups of free-ranging vervet monkeys. *American Zoologist* 22:519–529.
- Cheney DL, Seyfarth RM (2005) Constraints and preadaptations in the earliest stages of language evolution. *The Linguistic Review* 22:135–159.
- Cheney DL, Seyfarth RM (1996) Function and intention in the calls of non-human primates. *Proceedings of the British Academy*. pp 59–76
- Cheney DL, Seyfarth RM (1998) Why animals don't have language. *Tanner lectures on human values* 19:173–210.
- Cheney DL, Seyfarth RM (1985b) Social and non-social knowledge in vervet monkeys. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 187–201.
- Cheney DL, Seyfarth RM (1986) The recognition of social alliances by vervet monkeys. *Animal Behaviour* 34:1722–1731.
- Cheney DL, Seyfarth RM (1980) Vocal recognition in free-ranging vervet monkeys. *Animal Behaviour* 28:362–367. doi: 10.1016/S0003-3472(80)80044-3
- Cheney DL, Seyfarth RM (1992) *How monkeys see the world: inside the mind of another species*. University of Chicago Press, Chicago
- Cheney DL, Seyfarth RM (1982b) How vervet monkeys perceive their grunts: Field playback experiments. *Animal Behaviour* 30:739–751. doi: 10.1016/S0003-3472(82)80146-2
- Cheney DL, Seyfarth RM (1988) Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behaviour* 36:477–486. doi: 10.1016/S0003-3472(88)80018-6
- Cheney DL, Seyfarth RM, Palombit R (1996) The function and mechanisms underlying baboon "contact" barks. *Animal Behaviour* 52:507–518.
- Chomsky N (1959) Review of Verbal behaviour by B. F. Skinner. *Language* 35:26–58.
- Christiansen MH, Kirby S (2003) The hardest problem in science? In: Christiansen M, Kirby S (eds) *Language Evolution*. Oxford University Press, pp 1–15
- Clarke E, Reichard UH, Zuberbühler K (2006) The syntax and meaning of wild gibbon songs. *PLoS One* 1:e73. doi: 10.1371/journal.pone.0000073

- Clay Z, Smith CL, Blumstein DT (2012) Food-associated vocalizations in mammals and birds: what do these calls really mean? *Animal Behaviour* 83:323–330. doi: 10.1016/j.anbehav.2011.12.008
- Collins KT, McGreevy PD, Wheatley KE, Harcourt RG (2011) The influence of behavioural context on Weddell seal (*Leptonychotes weddellii*) airborne mother-pup vocalisation. *Behavioural Processes* 87:286–290. doi: 10.1016/j.beproc.2011.06.005
- Conner DA (1982) Geographic variation in short calls of pikas (*Ochotona princeps*). *Journal of Mammalogy* 48–52.
- Coss RG, McCowan B, Ramakrishnan U (2007) Threat-related acoustical differences in alarm calls by wild bonnet macaques (*Macaca radiata*) elicited by python and leopard models. *Ethology* 113:352–367. doi: 10.1111/j.1439-0310.2007.01336.x
- Crockford C, Herbinger I, Vigilant L, Boesch C (2004) Wild chimpanzees produce group-specific calls: a case for vocal learning? *Ethology* 110:221–243.
- Crockford C, Wittig RM, Seyfarth RM, Cheney DL (2007) Baboons eavesdrop to deduce mating opportunities. *Animal Behaviour* 73:885–890. doi: 10.1016/j.anbehav.2006.10.016
- Curio E (1978) The adaptive significance of avian mobbing: I. Teleonomic hypotheses and predictions. *Zeitschrift für Tierpsychologie* 48:175–183.
- Dandelot P (1959) Note sur la classification des cercopithèques du groupe aethiops. *Mammalia* 23:357–368.
- Darwin C (1872) *The Expression of the Emotions in Man and Animals*. John Murray, London
- Davidson SM, Wilkinson GS (2002) Geographic and individual variation in vocalizations by male *Saccopteryx bilineata* (Chiroptera: *Emballonuridae*). *Journal of Mammalogy* 83:526–535.
- Dawkins R, Krebs JR (1978) Animal signals: information or manipulation? In: Krebs JR, Davies, N. B. (eds) *Behavioural Ecology*. Blackwell, Oxford, pp 282–309
- Deacon TW (1997) *The symbolic species: the co-evolution of language and the brain*. W. W. Norton, New York
- Delgado RA (2007) Geographic variation in the long calls of male orangutans (*Pongo* spp.). *Ethology* 113:487–498. doi: 10.1111/j.1439-0310.2007.01345.x
- Dennett DC (1983) Intentional systems in cognitive ethology: The “Panglossian paradigm” defended. *Behavioral and Brain Sciences* 6:343–355. doi: 10.1017/S0140525X00016393
- Digweed SM, Rendall D (2009) Predator-associated vocalizations in North American red squirrels, *Tamiasciurus hudsonicus*: are alarm calls predator specific? *Animal Behaviour* 78:1135–1144.
- Doupe AJ, Kuhl PK (1999) Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience* 22:567–631.
- Dujardin E, Jürgens U (2006) Call type-specific differences in vocalization-related afferents to the periaqueductal gray of squirrel monkeys (*Saimiri sciureus*). *Behavioural Brain Research* 168:23–36. doi: 10.1016/j.bbr.2005.10.006

- Dujardin E, Jürgens U (2005) Afferents of vocalization-controlling periaqueductal regions in the squirrel monkey. *Brain Research* 1034:114–131. doi: 10.1016/j.brainres.2004.11.048
- Dunbar RIM (1974) Observations on the ecology and social organization of the green monkey, *Cercopithecus sabaues* in Senegal. *Primates* 15:341–350.
- Egnor SER, Hauser MD (2006) Noise-induced vocal modulation in cotton-top tamarins (*Saguinus oedipus*). *American Journal of Primatology* 68:1183–1190. doi: 10.1002/ajp.20317
- Egnor SER, Hauser MD (2004) A paradox in the evolution of primate vocal learning. *Trends in Neurosciences* 27:649–654.
- Egnor SER, Wickelgren JG, Hauser MD (2007) Tracking silence: adjusting vocal production to avoid acoustic interference. *Journal of Comparative Physiology A* 193:477–483. doi: 10.1007/s00359-006-0205-7
- Eiler KC, Banack SA (2004) Variability in the alarm call of golden-mantled ground squirrels (*Spermophilus lateralis* and *S. saturatus*). *Journal of Mammalogy* 85:43–50.
- Ellis J (2009) Anti-predator signals as advertisements: evidence in white-throated magpie-jays. *Ethology* 115:522–532.
- Ellsworth PC (2013) Appraisal theory: old and new questions. *Emotion Review* 5:125–131. doi: 10.1177/1754073912463617
- Ellsworth PC, Scherer KR (2003) Appraisal processes in emotion. In: Davidson RJ, Scherer KR, Goldsmith HH (eds) *Handbook of affective sciences*. Oxford University Press, Oxford, pp 572–595
- Elowson AM, Snowdon CT (1994) Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Animal Behaviour* 47:1267–1277.
- Elton S, Dunn J, Cardino A (2010) Size variation facilitates population divergence but does not explain it all: an example study from a widespread African monkey. *Biological Journal of the Linnean Society* 101:823–843.
- Engh AL, Hoffmeier RR, Cheney DL, Seyfarth RM (2006) Who, me? Can baboons infer the target of vocalizations? *Animal Behaviour* 71:381–387.
- Evans CS (1997) Referential signals. In: Owings DH, Beecher MD, Thompson NS (eds) *Perspectives in Ethology*. Plenum Press, New York, pp 99–143
- Evans CS, Evans L (2007) Representational signalling in birds. *Biology Letters* 3:8–11. doi: 10.1098/rsbl.2006.0561
- Evans CS, Evans L, Marler P (1993) On the meaning of alarm calls: Functional reference in an avian vocal system. *Animal Behaviour* 46:23–38.
- Evans CS, Marler P (1995) Language and animal communication: Parallels and contrasts. In: Roitblat HL, Arcady-Meyer J (eds) *Comparative approaches to cognitive science*. MIT Press, Cambridge MA, pp 341–382

- Evans CS, Marler P (1994) Food calling and audience effects in male chickens, *Gallus gallus*: their relationships to food availability, courtship and social facilitation. *Animal Behaviour* 47:1159–1170.
- Fairbanks LA (1980) Relationships among adult females in captive vervet monkeys: Testing a model of rank-related attractiveness. *Animal Behaviour* 28:853–859. doi: 10.1016/S0003-3472(80)80145-X
- Fairbanks LA, McGuire MT (1985) Relationships of vervet mothers with sons and daughters from one through three years of age. *Animal Behaviour* 33:40–50. doi: 10.1016/S0003-3472(85)80118-4
- Faragó T, Pongrácz P, Range F, et al. (2010) “The bone is mine”: affective and referential aspects of dog growls. *Animal Behaviour* 79:917–925. doi: 10.1016/j.anbehav.2010.01.005
- Fedigan L, Fedigan M (1988) *Cercopithecus aethiops*: a review of field studies. In: Gautier-Hion A, Bourlière F, Gautier J-P, Kingdon J (eds) *A Primate Radiation: Evolutionary Biology of the African Guenons*. Cambridge University Press, pp 389–411
- Fedurek P, Slocombe KE (2011) Primate vocal communication: a useful tool for understanding human speech and language evolution? *Human Biology* 83:153–173.
- Fichtel C, Hammerschmidt K (2002) Responses of redfronted lemurs (*Eulemur fulvus rufus*) to experimentally modified alarm calls: evidence for urgency-based changes in call structure. *Ethology* 108:763–777.
- Fichtel C, Hammerschmidt K, Jürgens U (2001) On the vocal expression of emotion. A multi-parametric analysis of different states of aversion in the squirrel monkey. *Behaviour* 138:97–116.
- Fichtel C, Kappeler PM (2002) Anti-predator behavior of group-living Malagasy primates: mixed evidence for a referential alarm call system. *Behavioral Ecology and Sociobiology* 262–275.
- Fichtel C, Perry S, Gros-Louis J (2005) Alarm calls of white-faced capuchin monkeys: an acoustic analysis. *Animal Behaviour* 70:165–176.
- Filmalter N (2010) A vegetation classification and management plan for the Hondekraal section of the Loskop Dam Nature Reserve. University of South Africa
- Fischer J (2011) Where is the information in animal communication? In: Menzel R, Fischer J (eds) *Animal Thinking: Contemporary Issues in Comparative Cognition*. The MIT Press, pp 151–161
- Fischer J (2003) Developmental modifications in the vocal behaviour of non-human primates. In: Ghazanfar A (ed) *Primate Audition*. CRC Press, pp 109–125
- Fischer J (2008) Transmission of acquired information in nonhuman primates. In: Byrne JH (ed) *Learning and memory: A comprehensive reference*. Elsevier, pp 299–313
- Fischer J (2006) Categorical perception. *Encyclopedia of Language and Linguistics*, 2nd ed. Elsevier, Oxford, pp 248–251

- Fischer J (2013) Information, inference and meaning in primate vocal behaviour. In: Stegmann U (ed) *Animal communication theory: information and influence*. Cambridge University Press, pp 297–319
- Fischer J (2010) Nothing to talk about. In: Frey UJ, Störmer C, Willführ KP (eds) *Homo Novus – A human without illusions*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 35–48
- Fischer J (1998) Barbary macaques categorize shrill barks into two call types. *Animal Behaviour* 55:799–807.
- Fischer J, Cheney DL, Seyfarth RM (2000) Development of infant baboons' responses to graded bark variants. *Proceedings of the Royal Society B: Biological Sciences* 267:2317–2321.
- 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 Behavior* 10:17–27. doi: 10.1111/j.1601-183X.2010.00610.x
- Fischer J, Hammerschmidt K (2001) Functional referents and acoustic similarity revisited: the case of Barbary macaque alarm calls. *Animal Cognition* 4:29–35.
- Fischer J, Hammerschmidt K, Cheney D, Seyfarth R (2001a) Acoustic features of female chacma baboon barks. *Ethology* 107:33–54.
- Fischer J, Hammerschmidt K, Cheney DL, Seyfarth RM (2002) Acoustic features of male baboon loud calls: Influences of context, age, and individuality. *Journal of the Acoustical Society of America* 111:1465–1474. doi: 10.1121/1.1433807
- Fischer J, Hammerschmidt K, Todt D (1998) Local variation in Barbary macaque shrill barks. *Animal Behaviour* 56:623–629.
- Fischer J, Hammerschmidt K, Todt D (1995) Factors affecting acoustic variation in Barbary macaque (*Macaca sylvanus*) disturbance calls. *Ethology* 101:51–66.
- Fischer J, Kitchen D, Seyfarth R, Cheney DL (2004) Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. *Behavioural Ecology and Sociobiology* 56:140–148.
- Fischer J, Metz M, Cheney D, Seyfarth R (2001b) Baboon responses to graded bark variants. *Animal Behaviour* 61:925–931.
- Fitch T (2010) *The evolution of language*. Cambridge University Press, Cambridge
- Fitch T (2003) Primate vocal production and its implications for auditory research. In: Ghazanfar A (ed) *Primate Audition. Ethology and Neurobiology*. CRC Press, pp 109–125
- Fitch T, Zuberbühler K (2013) Primate precursors to human language: Beyond discontinuity. In: Altenmüller E, Schmidt S, Zimmermann E (eds) *Evolution of emotional communication. From sounds in nonhuman mammals to speech and music in man*. Oxford University Press, Oxford, pp 26–49
- Fitch WT (2000) The evolution of speech: a comparative review. *Trends in Cognitive Sciences* 4:258–266.

- Fitch WT (2004) Kin selection and “mother tongues”: a neglected component in language evolution. In: Oller DK, Griebel U (eds) *Evolution of communication systems: A comparative approach*. MIT Press, Cambridge MA, pp 275–296
- Fitch WT, Hauser MD (2002) Unpacking “honesty”: vertebrate vocal production and the evolution of acoustic signals. *Springer handbook of auditory research* 16:65–137.
- Fitch WT, Hauser MD (1995) Vocal production in nonhuman primates: acoustics, physiology, and functional constraints on “honest” advertisement. *American Journal of Primatology* 37:
- Fox J, Weisberg S (2011) *An {R} Companion to Applied Regression*. Sage, Thousand Oaks CA
- Frederiksen P, Lawesson JE (1992) Vegetation types and patterns in Senegal based on multivariate analysis of field and NOAA-AVHRR satellite data. *Journal of Vegetation Science* 3:535–544.
- Freeberg TM, Dunbar RIM, Ord TJ (2012) Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367:1785–1801. doi: 10.1098/rstb.2011.0213
- Furrer RD, Manser MB (2009) Banded mongoose recruitment calls convey information about risk and not stimulus type. *Animal Behaviour* 78:195–201. doi: 10.1016/j.anbehav.2009.05.002
- Galat G, Galat-Luong A (1976) La colonisation de la mangrove par *Cercopithecus aethiops sabaesus* au Sénégal. *Terre et la Vie* 30:3–30.
- Gill SA, Sealy SG (2004) Functional reference in an alarm signal given during nest defence: see calls of yellow warblers denote brood-parasitic brown-headed cowbirds. *Behavioral Ecology and Sociobiology* 56:71–80. doi: 10.1007/s00265-003-0736-7
- Godfrey-Smith P (1991) Signal, decision, action. *The Journal of Philosophy* 88:709–722.
- Goudbeek M, Scherer K (2010) Beyond arousal: Valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America* 128:1322–1336. doi: 10.1121/1.3466853
- Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian Paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 205:581–598. doi: 10.1098/rspb.1979.0086
- Gouzoules S, Gouzoules H, Marler P (1984) Rhesus monkey (*Macaca mulatta*) screams: Representational signalling in the recruitment of agonistic aid. *Animal Behaviour* 32:182–193. doi: 10.1016/S0003-3472(84)80336-X
- Grafen A (1990) Biological signals as handicaps. *Journal of Theoretical Biology* 144:517–546.
- Greig EI, Pruett-Jones S (2010) Danger may enhance communication: predator calls alert females to male displays. *Behavioral Ecology* 21:1360–1366.
- Grice HP (1957) Meaning. *The Philosophical Review* 66:377. doi: 10.2307/2182440
- Groves C (2001) *Primate taxonomy*. Smithsonian Institution Press, Washington, DC

- Groves C (2005) Order primates. In: Wilson DE, Reeder DM (eds) *Mammal Species of the World: A Taxonomic and Geographic Reference*. JHU Press, pp 111–184
- Grubb P, Butynski TM, Oates JF, et al. (2003) Assessment of the diversity of African primates. *International Journal of Primatology* 24:1301–1357.
- Grüter C, Czaczkes TJ, Ratnieks FLW (2011) Decision making in ant foragers (*Lasius niger*) facing conflicting private and social information. *Behavioral Ecology and Sociobiology* 65:141–148. doi: 10.1007/s00265-010-1020-2
- Grüter C, Ratnieks FLW (2011) Honeybee foragers increase the use of waggle dance information when private information becomes unrewarding. *Animal Behaviour*
- Guilford T, Dawkins MS (1991) Receiver psychology and the evolution of animal signals. *Animal Behaviour* 42:1–14.
- Gyger M, Karakashian SJ, Marler P (1986) Avian alarm calling: is there an audience effect? *Animal Behaviour* 34:1570–1572.
- Hage SR (2010) Neuronal networks involved in the generation of vocalization. In: Stefan M. Brudzynski (ed) *Handbook of Behavioral Neuroscience*. Elsevier, pp 339–349
- Hage SR (2013) Audio–vocal interactions during vocal communication in squirrel monkeys and their neurobiological implications. *Journal of Comparative Physiology A* 199:663–668. doi: 10.1007/s00359-013-0810-1
- Hage SR, Gavrilov N, Nieder A (2013) Cognitive control of distinct vocalizations in rhesus monkeys. *Journal of Cognitive Neuroscience* 1–10. doi: 10.1162/jocn\_a\_00428
- Hamilton WD (1964) The genetical evolution of social behaviour. I. *Journal of Theoretical Biology* 7:1–16.
- Hamilton WD (1963) The evolution of altruistic behavior. *American Naturalist* 97:354–356.
- Hammerschmidt K, Fischer J (2008) Constraints in primate vocal production. In: Oller DK, Griebel U (eds) *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication*. MIT Press, pp 93–121
- Hammerschmidt K, Fischer J (1998a) Maternal discrimination of offspring vocalizations in Barbary macaques (*Macaca sylvanus*). *Primates* 39:231–236.
- Hammerschmidt K, Fischer J (1998b) The vocal repertoire of Barbary macaques: A quantitative analysis of a graded signal system. *Ethology* 104:203–216.
- Hammerschmidt K, Jürgens U, Freudenstein T (2001) Vocal development in squirrel monkeys. *Behaviour* 138:1179–1204. doi: 10.1163/156853901753287190
- Hammerschmidt K, Newman JD, Champoux M, Suomi SJ (2000) Changes in rhesus macaque “coo” vocalizations during early development. *Ethology* 106:873–886.
- Harley TA (2001) *The psychology of language: from data to theory*, 2nd ed. Psychology Press
- Harnad SR (1987) *Categorical perception: The groundwork of cognition*. Cambridge University Press

- Hasson O (1991) Pursuit-deterrent signals: Communication between prey and predator. *Trends in Ecology & Evolution* 6:325–329.
- Haus T, Akom E, Agwanda B, et al. (2013) Mitochondrial diversity and distribution of African green monkeys (*Chlorocebus Gray, 1870*). *American Journal of Primatology* 75:350–360. doi: 10.1002/ajp.22113
- Hauser M (1997) *The evolution of communication*, New edition. MIT Press, Cambridge, Massachusetts
- Hauser M, Fitch T (2003) What are the uniquely human components of the language faculty? In: Christiansen M, Kirby S (eds) *Language Evolution*. Oxford University Press, pp 158–181
- Hauser MD (1993a) The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight, and social context. *The American Naturalist* 142:528–542.
- Hauser MD (1998) Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Animal Behaviour* 55:1647–1658.
- Hauser MD (1993b) Do vervet monkey infants cry wolf? *Animal Behaviour* 45:1242–1244. doi: 10.1006/anbe.1993.1148
- Hauser MD, Chomsky N, Fitch WT (2002) The faculty of language: What is it, who has it, and how did it evolve? *Science* 298:1569–1579. doi: 10.1126/science.298.5598.1569
- Henzi SP, Forshaw N, Boner R, et al. (2013) Scalar social dynamics in female vervet monkey cohorts. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368:20120351–20120351. doi: 10.1098/rstb.2012.0351
- Heyes C (2012) Simple minds: a qualified defence of associative learning. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367:2695–2703. doi: 10.1098/rstb.2012.0217
- Hill WCO (1966) *Primates comparative anatomy and taxonomy VI Catarrhini Cercopithecoidea Cercopithecinae*. Edinburgh University Press, Edinburgh
- Hinde R (1981) Animal signals: Ethological and games-theory approaches are not incompatible. *Animal Behaviour* 29:535–542. doi: 10.1016/S0003-3472(81)80116-9
- Hobin JA, Goosens KA, Maren S (2003) Context-dependent neuronal activity in the lateral amygdala represents fear memories after extinction. *The Journal of Neuroscience* 23:8410–8416.
- Hockett C (1960) The origin of speech. *Scientific American* 203:88–96.
- Hockett C, Altmann S (1968) A note on design features. In: Sebeok TA (ed) *Animal communication: Techniques of study and results of research*. Indiana University Press, pp 61–72
- Huff NC, Hernandez JA, Fecteau ME, et al. (2011) Revealing context-specific conditioned fear memories with full immersion virtual reality. *Frontiers in Behavioral Neuroscience*. doi: 10.3389/fnbeh.2011.00075

- 
- Hurford JR (2003) The language mosaic and its evolution. In: Christiansen MH, Kirby S (eds) Language Evolution. Oxford University Press, Oxford, pp 38–57
- Hurford JR (2007) The origins of meaning. Oxford University Press, Oxford
- Jackendoff R (1999) Possible stages in the evolution of the language capacity. Trends in Cognitive Sciences 3:272–279.
- Janik VM, Slater PJB (2000) The different roles of social learning in vocal communication. Animal Behaviour 60:1–11.
- Janik VM, Slater PJB (1997) Vocal learning in mammals. Advances in the Study of Behavior: Volume 26 59.
- Jarvis ED (2006) Selection for and against vocal learning in birds and mammals. Ornithological Science 5:5–14.
- Jespersen O (1922) Language; its nature, development and origin. London : G. Allen & Unwin, Ltd.
- Jovanovic T, Gouzoules H (2001) Effects of nonmaternal restraint on the vocalizations of infant rhesus monkeys (*Macaca mulatta*). American Journal of Primatology 53:33–45. doi: 10.1002/1098-2345(200101)53:1<33::AID-AJP3>3.0.CO;2-C
- Jürgens U (1976a) Projections from the cortical larynx area in the squirrel monkey. Experimental Brain Research 25:401–411.
- Jürgens U (2009) The neural control of vocalization in mammals: a review. Journal of Voice 23:1–10.
- Jürgens U (1998) Neuronal control of mammalian vocalization, with special reference to the squirrel monkey. Naturwissenschaften 85:376–388.
- Jürgens U (1976b) Reinforcing concomitants of electrically elicited vocalizations. Experimental Brain Research 26:203–214. doi: 10.1007/BF00238284
- Jürgens U (1979) Vocalization as an emotional indicator a neuroethological study in the squirrel monkey. Behaviour 88–117.
- Jürgens U (1992) On the neurobiology of vocal communication. In: Papoušek H, Jürgens U, Papoušek M (eds) Nonverbal Vocal Communication: Comparative and Developmental Approaches. Cambridge University Press, pp 31–42
- Kaufman L, Rousseeuw PJ (2005) Finding groups in data: An introduction to cluster analysis, 2nd Revised edition. Smithsonian Institution Press, Washington, DC
- Keenan S, Lemasson A, Zuberbühler K (2013) Graded or discrete? A quantitative analysis of Campbell's monkey alarm calls. Animal Behaviour 85:109–118. doi: 10.1016/j.anbehav.2012.10.014
- Kingdon J (1997) The Kingdon field guide to African mammals. Academic Press, London
- Kirby S (1999) Function, selection, and innateness: The emergence of language universals. Oxford University Press
-

- Kirchhof J, Hammerschmidt K (2006) Functionally referential alarm calls in tamarins (*Saguinus fuscicollis* and *Saguinus mystax*) - evidence from playback experiments. *Ethology* 112:346–354. doi: 10.1111/j.1439-0310.2006.01165.x
- Kiriazis J, Slobodchikoff CN (2006) Perceptual specificity in the alarm calls of Gunnison's prairie dogs. *Behavioural Processes* 73:29–35. doi: 10.1016/j.beproc.2006.01.015
- Kitchen DM, Seyfarth RM, Fischer J, Cheney DL (2003) Loud calls as indicators of dominance in male baboons (*Papio cynocephalus ursinus*). *Behavioral Ecology and Sociobiology* 53:374–384.
- Kitzmann CD, Caine NG (2009) Marmoset (*Callithrix geoffroyi*) Food-Associated Calls are Functionally Referential. *Ethology* 115:439–448. doi: 10.1111/j.1439-0310.2009.01622.x
- Koda H, Oyakawa C, Kato A, Masataka (2007) Experimental evidence for the volitional control of vocal production in an immature gibbon. *Behaviour* 144:681–692. doi: 10.1163/156853907781347817
- Krebs JR, Dawkins R (1984) Animal signals: mind-reading and manipulation. *Behavioural Ecology: an evolutionary approach* 2:380–402.
- Kuhl PK, Miller JD (1975) Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* 190:69.
- Kuypers HG (1958) Corticobulbar connexions to the pons and lower brain-stem in man. *Brain* 81:364–388.
- Lameira AR, Delgado RA, Wich SA (2010) Review of geographic variation in terrestrial mammalian acoustic signals: Human speech variation in a comparative perspective. *Journal of Evolutionary Psychology* 8:309–332. doi: 10.1556/JEP.8.2010.4.2
- Lancaster JB (1975) *Primate behavior and the emergence of human culture*. Holt, Rinehart and Winston
- Lang PJ, Bradley MM, Cuthbert BN (1998) Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biological Psychiatry* 44:1248–1263.
- Langmore NE, Mulder RA (1992) A novel context for bird song: Predator calls prompt male singing in the kleptogamous superb fairy-wren, *Malurus cyaneus*. *Ethology* 90:143–153.
- Lazarus RS (1966) *Psychological stress and the coping process*. McGraw-Hill, New York, NY, US
- Lazarus RS (1991) *Emotion and adaptation*. Oxford University Press, New York
- Leander JD, Milan MA, Jasper KB, Heaton KL (1972) Schedule control of the vocal behavior of Cebus monkeys. *Journal of the Experimental Analysis of Behavior* 17:229–235. doi: 10.1901/jeab.1972.17-229
- Lernould J (1988) Classification and geographical distribution of guenons: a review. In: Gautier-Hion A, Bourlière F, Gautier JP, Kingdon J (eds) *Cambridge University Press*, pp 54–78
- Leventhal H, Scherer K (1987) The relationship of emotion to cognition: A functional approach to a semantic controversy. *Cognition and Emotion* 1:3–28.

- Liberman AM (1957) Some results of research on speech perception. *The Journal of the Acoustical Society of America* 29:117–123. doi: 10.1121/1.1908635
- Liberman AM, Pisoni DB (1977) Evidence for a special speech-perceiving subsystem in the human. *Recognition of complex acoustic signals*. Berlin: Dahlem Konferenzen
- Lieblich AK, Symmes D, Newman JD, Shapiro M (1980) Development of the isolation peep in laboratory-bred squirrel monkeys. *Animal Behaviour* 28:1–9.
- Lotem A, Halpern JY (2012) Coevolution of learning and data-acquisition mechanisms: a model for cognitive evolution. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 367:2686–2694. doi: 10.1098/rstb.2012.0213
- Macedonia JM (1990) What is communicated in the antipredator calls of lemurs: Evidence from playback experiments with ringtailed and ruffed lemurs. *Ethology* 86:177–190. doi: 10.1111/j.1439-0310.1990.tb00428.x
- Macedonia JM, Evans CS (1993) Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology* 93:177–177.
- Maciej P, Patzelt A, Ndao I, et al. (2013) Social monitoring in a multilevel society: a playback study with male Guinea baboons. *Behavioral Ecology and Sociobiology* 67:61–68. doi: 10.1007/s00265-012-1425-1
- Malle BF (2002) The relation between language and theory of mind in development and evolution. In: Givón T, Malle BF (eds) *The evolution of language out of pre-language*. Benjamins, Amsterdam, pp 265–284.
- Manser MB (2001) The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B: Biological Sciences* 268:2315–2324.
- Manser MB, Bell MB, Fletcher LB (2001) The information that receivers extract from alarm calls in suricates. *Proceedings: Biological Sciences* 286:2485–2491.
- Manser MB, Seyfarth RM, Cheney DL (2002) Suricate alarm calls signal predator class and urgency. *Trends in Cognitive Sciences* 6:55–57.
- Marchant JN, Whittaker X, Broom DM (2001) Vocalisations of the adult female domestic pig during a standard human approach test and their relationships with behavioural and heart rate measures. *Applied Animal Behaviour Science* 72:23–39.
- Marler P (1975) On the origin of speech from animal sounds. In: Kavanagh JF, Cutting J (eds) *The role of speech in language*. MIT Press, Cambridge MA, pp 11–37
- Marler P (1976) Social organisation, communications and graded signals: the chimpanzee and the gorilla. *Growing points in Ethology*. Cambridge University Press, Cambridge, pp 239–280
- Marler P (1977) Primate vocalization: Affective or symbolic? In: Bourne G (ed) *Progress in ape research*. Academic Press, New York, pp 85–96
- Marler P (1955) Characteristics of some animal calls. *Nature* 176:6–8. doi:10.1038/176006a0

- Marler P (1968) Aggregation and dispersal: Two functions in primate communication. In: Jay P (ed) *Primates: Studies in adaptation and variability*. Holt, Rinehart and Winston, New York, pp 420-438
- Marler P, Evans CS, Hauser MD (1992) Animal signals: motivational, referential, or both? In: Papoušek H, Jürgens U, Papoušek M (eds) *Nonverbal vocal communication: comparative and developmental approaches*. Cambridge University Press, Cambridge, pp 66–86
- Marshall AJ, Wrangham RW, Arcadi AC (1999) Does learning affect the structure of vocalizations in chimpanzees? *Animal Behaviour* 58:825–830.
- Maynard Smith J (1978) Optimization theory in evolution. *Annual Review of Ecology and Systematics* 9:31–56.
- Maynard Smith J (1985) Sexual selection, handicaps and true fitness. *Journal of Theoretical Biology* 115:1–8. doi: 10.1016/S0022-5193(85)80003-5
- Maynard Smith J (1991) Honest signalling: the Philip Sidney game. *Animal Behaviour* 42:1034–1035.
- Maynard Smith J, Harper DGC (1995) Animal signals: models and terminology. *Journal of Theoretical Biology* 177:305–311.
- Maynard Smith J, Szathmáry E (1995) *The major transitions in evolution*. W. H. Freeman, Oxford
- McComb K, Semple S (2005) Coevolution of vocal communication and sociality in primates. *Biology Letters* 1:381–385.
- 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. *The Journal of the Acoustical Society of America* 129:1631. doi: 10.1121/1.3531944
- Méndez-Cárdenas M, Randrianambinina B, Rabesandratana A, et al. (2008) Geographic variation in loud calls of sportive lemurs (*Lepilemur* spp.) and their implications for conservation. *American Journal of Primatology* 70:
- Mendl M, Burman OHP, Parker RMA, Paul ES (2009) Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science* 118:161–181. doi: 10.1016/j.applanim.2009.02.023
- Mendl M, Burman OHP, Paul ES (2010) An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences* 277:2895–2904. doi: 10.1098/rspb.2010.0303
- Meyer D, Hodges JK, Rinaldi D, et al. (2012) Acoustic structure of male loud-calls support molecular phylogeny of Sumatran and Javanese leaf monkeys (genus *Presbytis*). *BMC Evolutionary Biology* 12:16.
- Meyer D, Rinaldi ID, Ramlee H, et al. (2011) Mitochondrial phylogeny of leaf monkeys (genus *Presbytis*, *Eschscholtz*, 1821) with implications for taxonomy and conservation. *Molecular Phylogenetics and Evolution* 59:311–319. doi: 10.1016/j.ympev.2011.02.015

- Moors A, Ellsworth PC, Scherer KR, Frijda NH (2013) Appraisal theories of emotion: State of the art and future development. *Emotion Review* 5:119–124. doi: 10.1177/1754073912468165
- Morse PA, Snowden CT (1975) An investigation of categorical speech discrimination by rhesus monkeys. *Perception & Psychophysics* 17:9–16. doi: 10.3758/BF03203991
- Morton ES (1977) On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist* 111:855–869.
- Morton ES (1982) Grading, discreteness, redundancy, and motivation-structural rules. In: Kroodsma DE, Miller EH (eds) *Acoustic Communication in Birds: Production, perception, and design features of sounds*. Academic Press, New York, p 183
- Morton ES (1975) Ecological Sources of Selection on Avian Sounds. *The American Naturalist* 109:17. doi: 10.1086/282971
- Mundry R, Sommer C (2007) Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour* 74:965–976. doi: 10.1016/j.anbehav.2006.12.028
- Murphy D, Lea SEG, Zuberbühler K (2013) Male blue monkey alarm calls encode predator type and distance. *Animal Behaviour* 85:119–125. doi: 10.1016/j.anbehav.2012.10.015
- Naguib M, Hammerschmidt K, Wirth J (2001) Microgeographic variation, habitat effects and individual signature cues in calls of chiffchaffs *Phylloscopus collybita canarensis*. *Ethology* 107:341–355. doi: 10.1046/j.1439-0310.2001.00669.x
- Napier PH (1981) Catalogue of primates in the British Museum (Natural History) and elsewhere in the British Isles. Part II: Family *Cercopithecidae*, subfamily *Cercopithecinae*. British Museum (Natural History), London.
- Nelson DA, Marler P (1989) Categorical perception of a natural stimulus continuum: birdsong. *Science* 244:976–978.
- Neumann C, Assahad G, Hammerschmidt K, et al. (2010) Loud calls in male crested macaques, *Macaca nigra*: a signal of dominance in a tolerant species. *Animal Behaviour* 79:187–193. doi: 10.1016/j.anbehav.2009.10.026
- Nietsch A, Kopp ML (1998) Role of vocalization in species differentiation of Sulawesi tarsiers. *Folia Primatologica* 69:371–378.
- Norcross JL, Newman JD (1999) Effects of separation and novelty on distress vocalizations and cortisol in the common marmoset (*Callithrix jacchus*). *American Journal of Primatology* 47:209–222. doi: 10.1002/(SICI)1098-2345(1999)47:3<209::AID-AJP3>3.0.CO;2-0
- Normansell L, Panksepp J (2011) Glutamatergic modulation of separation distress: profound emotional effects of excitatory amino acids in chicks. *Neuroscience and Biobehavioral Reviews* 35:1890–1901. doi: 10.1016/j.neubiorev.2011.06.004
- Nottebohm F (1972) The origins of vocal learning. *American Naturalist* 116–140.
- Oates JF, Trocco TF (1983) Taxonomy and phylogeny of black-and-white colobus monkeys inferences from an analysis of loud call variation. *Folia Primatologica* 40:83–113.

- Oller DK, Griebel U (2008) Signal and functional flexibility in the emergence of communication systems: The editors introduction. In: Oller DK, Griebel U (eds) *Evolution of Communicative Flexibility: Complexity, Creativity, and Adaptability in Human and Animal Communication*. Mit Press, pp 3–8
- Owings DH, Hennessy DF (1984) The importance of variation in sciurid visual and vocal communication. In: Murie JO, Michener GR (eds) *The biology of ground dwelling squirrels*. University of Nebraska Press, Omaha, pp 171–200
- Owren MJ (1990a) Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*). 1. Natural Calls. *Journal of Comparative Psychology* 104:20–28.
- Owren MJ (1990b) Acoustic classification of alarm calls by vervet monkeys (*Cercopithecus aethiops*) and humans (*Homo sapiens*): II. Synthetic calls. *Journal of Comparative Psychology* 104:29–40.
- Owren MJ, Amoss RT, Rendall D (2011) Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology* 73:530–544. doi: 10.1002/ajp.20913
- Owren MJ, Bernacki RH (1988) The acoustic features of vervet monkey alarm calls. *The Journal of the Acoustical Society of America* 83:1927–1935.
- Owren MJ, Dieter JA, Seyfarth RM, Cheney DL (1992) “Food” calls produced by adult female rhesus (*Macaca Mulatta*) and Japanese (*M. Fuscata*) macaques, their normally-raised offspring, and offspring cross-fostered between species. *Behaviour* 120:218–231. doi: 10.1163/156853992X00615
- Owren MJ, Rendall D (1997) An affect-conditioning model of nonhuman primate. In: Owings DH, Beecher MD, Thompson NS (eds) *Perspectives in Ethology: Communication: Vol.12*. Springer, Worcester, Massachusetts, pp 299–346
- Owren MJ, Rendall D (2001) Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. *Evolutionary Anthropology* 10:58–71.
- Panksepp J (2011) The basic emotional circuits of mammalian brains: Do animals have affective lives? *Neuroscience & Biobehavioral Reviews* 35:1791–1804. doi: 10.1016/j.neubiorev.2011.08.003
- Partan S, Marler P (1999) Communication goes multimodal. *Science* 283:1272–3.
- Pasternak G, Brown LR, Kienzle S, et al. (2013) Population ecology of vervet monkeys in a high latitude, semi-arid riparian woodland. *Koedoe-African Protected Area Conservation and Science*. doi: 10.4102/koedoe.v55i1.1078
- Pearce JM, Bouton ME (2001) Theories of associative learning in animals. *Annual Review of Psychology* 52:111–139.
- Peirce CS (1958) *The collected papers*. Harvard University Press

- Penn DC, Povinelli DJ (2007) On the lack of evidence that non-human animals possess anything remotely resembling a “theory of mind.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 362:731–744. doi: 10.1098/rstb.2006.2023
- Peterson GE, Barney HL (1952) Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America* 24:175–184. doi: 10.1121/1.1906875
- 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. doi: 10.1016/j.anbehav.2005.08.021
- Pinker S, Bloom P (1990) Natural language and natural selection. *Behavioral and Brain Sciences* 13:707–727. doi: 10.1017/S0140525X00081061
- Poole JH, Tyack PL, Stoeger-Horwath AS, Watwood S (2005) Animal behaviour: Elephants are capable of vocal learning. *Nature* 434:455–456. doi: 10.1038/434455a
- Premack D (1975) On the origins of language. In: Gazzaniga MS, Blakemore C (eds) *Handbook of Psychobiology*. Academic Press, New York, pp 591–605
- Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1:515–526. doi: 10.1017/S0140525X00076512
- Price T, Fischer J (2013) Meaning attribution in the West African green monkey: influence of call type and context. *Animal Cognition*. doi: DOI 10.1007/s10071-013-0660-9
- R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria
- Reby D, McComb K (2003) Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. *Animal Behaviour* 65:519–530.
- Rendall D, Cheney DL, Seyfarth RM (2000) Proximate factors mediating “contact” calls in adult female baboons (*Papio cynocephalus ursinus*) and their infants. *Journal of Comparative Psychology* 114:36–46.
- Rendall D, Owren MJ, Ryan MJ (2009) What do animal signals mean? *Animal Behaviour* 78:233–240. doi: 10.1016/j.anbehav.2009.06.007
- Rendall D, Seyfarth RM, Cheney DL, Owren MJ (1999) The meaning and function of grunt variants in baboons. *Animal Behaviour* 57:583–592.
- Rescorla RA (1988) Behavioral studies of Pavlovian conditioning. *Annual Review of Neuroscience* 11:329–352.
- Ridley M (2004) *Evolution*. Blackwell Science Ltd, Malden, MA
- Riede T, Herzog H, Hammerschmidt K, et al. (2001) The harmonic-to-noise ratio applied to dog barks. *The Journal of the Acoustical Society of America* 110:2191. doi: 10.1121/1.1398052
- Robinson SR (1980) Antipredator behaviour and predator recognition in Belding’s ground squirrels. *Animal Behaviour* 28:840–852.

- Roseman IJ, Smith CA (2001) Appraisal theory: Overview, assumptions, varieties, controversies. In: Scherer KR, Schorr A, Johnstone T (eds) *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press, New York, NY, US, pp 3–19
- Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* 20:53–65.
- Le Roux A, Cherry MI, Manser MB (2008) The audience effect in a facultatively social mammal, the yellow mongoose, *Cynictis penicillata*. *Animal Behaviour* 75:943–949.
- Rowe C (1999) Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour* 58:921–931. doi: 10.1006/anbe.1999.1242
- Rowell TE, Hinde RA (1962) Vocal Communication by the Rhesus (*Macaca mulatta*). *Proceedings of the Zoological Society of London* 138:279–294. doi: 10.1111/j.1469-7998.1962.tb05698.x
- Scarantino A (2010) Animal communication between information and influence. *Animal Behaviour* 30:1e5.
- Scarantino A (in press) Rethinking functional reference. *Philosophy of Science*
- Scherer KR (1989) Vocal correlates of emotional arousal and affective disturbance. *Handbook of social psychophysiology* 165–197.
- Scherer KR (2009a) Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364:3459–3474. doi: 10.1098/rstb.2009.0141
- Scherer KR (2003) Vocal communication of emotion: A review of research paradigms. *Speech Communication* 40:227–256.
- Scherer KR (2013) The nature and dynamics of relevance and valence appraisals: Theoretical advances and recent evidence. *Emotion Review* 5:150–162. doi: 10.1177/1754073912468166
- Scherer KR (2001) Appraisal considered as a process of multilevel sequential checking. *Appraisal Processes in Emotion: Theory, Methods, Research* 92:120.
- Scherer KR (2009b) The dynamic architecture of emotion: Evidence for the component process model. *Cognition and Emotion* 23:1307–1351. doi: 10.1080/02699930902928969
- Scherer KR, Johnstone T, Klasmeyer G (2003) Vocal expression of emotion. *Handbook of affective sciences* 433–456.
- Schneider C, Hodges K, Fischer J, Hammerschmidt K (2008) Acoustic niches of Siberut primates. *International Journal of Primatology* 29:601–613.
- Schneiderová I, Policht R (2011) Alarm calls of the European ground squirrel *Spermophilus citellus* and the Taurus ground squirrel *S. taurensis* encode information about caller identity. *Bioacoustics* 20:29–43.

- Schoenemann PT (2005) Conceptual complexity and the brain: Understanding language origins. In: Minett JW, Wang WS (eds) *Language acquisition, change and emergence: Essays in evolutionary linguistics*. City University of Hong Kong Press, pp 47–94
- Schrader L, Hammerschmidt K (1997) Computer aided analysis of acoustic parameters in animal vocalisations. A multiparametric approach. *Bioacoustics* 7:247–265.
- Schrader L, Todt D (1998) Vocal quality is correlated with levels of stress hormones in domestic pigs. *Ethology* 104:859–876.
- Scott-Phillips TC (2009) Animal communication: insights from linguistic pragmatics. *Animal Behaviour* 79:e1–e4. doi: 10.1016/j.anbehav.2009.10.013
- Scott-Phillips TC (2008) Defining biological communication. *Journal of Evolutionary Biology* 21:387–395.
- Searcy WA, Nowicki S (2010) *The evolution of animal communication: Reliability and deception in signaling systems*. Princeton University Press
- Semple S (2001) Individuality and male discrimination of female copulation calls in the yellow baboon. *Animal Behaviour* 61:1023–1028. doi: 10.1006/anbe.2001.1692
- Semple S, Higham JP (2013) Primate signals: Current issues and perspectives. *American Journal of Primatology* 75:613–620. doi: 10.1002/ajp.22139
- Semple S, McComb K, Alberts S, Altmann J (2002) Information content of female copulation calls in yellow baboons. *American Journal of Primatology* 56:43–56.
- Servedio MR, Doorn GSV, Kopp M, et al. (2011) Magic traits in speciation: “magic” but not rare? *Trends in Ecology & Evolution* 26:389–397. doi: 10.1016/j.tree.2011.04.005
- Seyfarth R, Cheney D (1990) The assessment by vervet monkeys of their own and another species’ alarm calls. *Animal Behaviour* 40:754–764. doi: 10.1016/S0003-3472(05)80704-3
- Seyfarth RM (1980) The distribution of grooming and related behaviours among adult female vervet monkeys. *Animal Behaviour* 28:798–813. doi: 10.1016/S0003-3472(80)80140-0
- Seyfarth RM, Cheney DL (2003a) Meaning and emotion in animal vocalizations. *Annals of the New York Academy of Sciences* 1000:32–55.
- Seyfarth RM, Cheney DL (1980) The ontogeny of vervet monkey alarm-calling behavior: A preliminary report. *Zeitschrift für Tierpsychologie* 54:37–56.
- Seyfarth RM, Cheney DL (2003b) Signalers and receivers in animal communication. *Annual Review of Psychology* 54:145–173.
- Seyfarth RM, Cheney DL (1997) Some general features of vocal development in nonhuman primates. In: Snowdon CT, Hausberger M (eds) *Social influences on vocal development*. Cambridge University Press, pp 249–273
- Seyfarth RM, Cheney DL (2010) Production, usage, and comprehension in animal vocalizations. *Brain and Language* 115:92–100. doi: 10.1016/j.bandl.2009.10.003

- Seyfarth RM, Cheney DL (1986) Vocal development in vervet monkeys. *Animal Behaviour* 34:1640–1658. doi: 10.1016/S0003-3472(86)80252-4
- Seyfarth RM, Cheney DL (1993) Meaning, reference, and intentionality in the natural vocalizations of monkeys. *Language and communication: Comparative Perspectives* 195–220.
- Seyfarth RM, Cheney DL, Bergman T, et al. (2010) The central importance of information in studies of animal communication. *Animal Behaviour* 80:3–8. doi: 10.1016/j.anbehav.2010.04.012
- Seyfarth RM, Cheney DL, Bergman TJ (2005) Primate social cognition and the origins of language. *Trends in Cognitive Sciences* 9:264–266.
- Seyfarth RM, Cheney DL, Marler P (1980a) Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210:801–803.
- Seyfarth RM, Cheney DL, Marler P (1980b) Vervet monkey alarm calls. Semantic communication in a free-ranging primate. *Animal Behaviour* 28:1070–1094. doi: 10.1016/S0003-3472(80)80097-2
- Sherman PW (1977) Nepotism and the evolution of alarm calls. *Science* 197:1246–1253.
- Silk JB, Kaldor E, Boyd R (2000) Cheap talk when interests conflict. *Animal Behaviour* 59:423–432. doi: 10.1006/anbe.1999.1312
- Slabbekoorn H, Smith TB (2002) Bird song, ecology and speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 357:493–503. doi: 10.1098/rstb.2001.1056
- Slater P, Lester N (1982) Minimising errors in splitting behaviour into bouts. *Behaviour* 153–161.
- Slobodchikoff CN (2002) Cognition and communication in prairie dogs. *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition* 257–264.
- Slobodchikoff CN, Ackers SH, Van Ert M (1998) Geographic variation in alarm calls of Gunnison's prairie dogs. *Journal of Mammalogy* 1265–1272.
- Slobodchikoff CN, Coast R (1980) Dialects in the alarm calls of prairie dogs. *Behavioral Ecology and Sociobiology* 7:49–53.
- Slocombe K, Zuberbühler K (2005) Functionally referential communication in a chimpanzee. *Current Biology* 15:1779–1784. doi: 10.1016/j.cub.2005.08.068
- Slocombe KE, Zuberbühler K (2007) Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences* 104:17228.
- Smith MJ, Hunter D (2005) Temporal and geographic variation in the advertisement call of the booroolong frog (*Litoria booroolongensis*: Anura: *Hylidae*). *Ethology* 111:1103–1115.
- Smith WJ (1977) *The behavior of communicating: an ethological approach*. Harvard University Press, Cambridge, Mass. and London, England

- Snowdon CT (2004) Social processes in the evolution of complex cognition and communication. In: Oller DK, Griebel U (eds) *Evolution of communication systems: a comparative approach*. MIT Press, Cambridge MA, pp 131–150
- Snowdon CT (1990) Language capacities of nonhuman animals. *American Journal of Physical Anthropology* 33:215–243.
- Snowdon CT, Elowson AM (1999) Pygmy marmosets modify call structure when paired. *Ethology* 105:893–908.
- Snowdon CT, Hodun A, Rosenberger AL, Coimbra-Filho AF (1986) Long-call structure and its relation to taxonomy in lion tamarins. *American Journal of Primatology* 11:253–261.
- Soltis J, Blowers TE, Savage A (2011) Measuring positive and negative affect in the voiced sounds of African elephants (*Loxodonta africana*). *Journal of the Acoustical Society of America* 129:1059–1066. doi: 10.1121/1.3531798
- Soltis J, Leong K, Savage A (2005) African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers. *Animal Behaviour* 70:589–599. doi: 10.1016/j.anbehav.2004.11.016
- Stoeger AS, Mietchen D, Oh S, et al. (2012) An Asian Elephant Imitates Human Speech. *Current Biology* 22:2144–2148. doi: 10.1016/j.cub.2012.09.022
- Struhsaker TT (1970) Phylogenetic implications of some vocalizations of *Cercopithecus* monkeys. In: Napier JR, Napier PH (eds) *Old World monkeys: evolution, systematics, and behavior*. Academic Press, New York, pp 365–444
- Struhsaker TT (1967a) Ecology of Vervet Monkeys (*Cercopithecus Aethiops*) in The Masai-Amboseli Game Reserve, Kenya. *Ecology* 48:891–904.
- Struhsaker TT (2010) *Vocalizations. The Red Colobus Monkeys: Variation in Demography, Behavior, and Ecology of Endangered Species*. Oxford University Press, pp 18–44
- Struhsaker TT (1967b) Auditory communication among vervet monkeys (*Cercopithecus aethiops*). In: Altmann S (ed) *Social communication among primates*. University of Chicago Press, Chicago, pp 281–325
- Struhsaker TT (1967c) Social structure among vervet monkeys (*Cercopithecus aethiops*). *Behaviour* 83–121.
- Studdert-Kennedy M (2005) How did language go discrete? In: Tallerman M (ed) *Language Origins. Perspectives on evolution*. Oxford University Press, Oxford, pp 48–67
- Sutton D, Larson C, Lindeman RC (1974) Neocortical and limbic lesion effects on primate phonation. *Brain Research* 71:61–75.
- Sutton D, Larson C, Taylor EM, Lindeman RC (1973) Vocalization in rhesus monkeys: Conditionability. *Brain Research* 52:225–231.
- Svendsen GE (1979) Territoriality and behavior in a population of pikas (*Ochotona princeps*). *Journal of Mammalogy* 324–330.

- Taylor AM, Reby D, McComb K (2009) Context-related variation in the vocal growling behaviour of the domestic dog (*Canis familiaris*). *Ethology* 115:905–915. doi: 10.1111/j.1439-0310.2009.01681.x
- Theis KR, Greene KM, Benson-Amram SR, Holekamp KE (2007) Sources of variation in the long-distance vocalizations of spotted hyenas. *Behaviour* 144:557–584.
- Thinh V, Hallam C, Roos C, Hammerschmidt K (2011) Concordance between vocal and genetic diversity in crested gibbons. *BMC Evolutionary Biology* 11:36.
- Thinh VN, Mootnick AR, Geissmann T, et al. (2010) Mitochondrial evidence for multiple radiations in the evolutionary history of small apes. *BMC Evolutionary Biology* 10:74. doi: 10.1186/1471-2148-10-74
- Tibbetts EA, Dale J (2007) Individual recognition: it is good to be different. *Trends in Ecology & Evolution* 22:529–537.
- Tincoff R, Hauser MD (2009) Cognitive basis for language evolution in nonhuman primates. *Encyclopedia of language and linguistics*
- Todt D (1986) Hinweischarakter und mittlerfunktion von verhalten. *Zt. Semiotik* 8:183–232.
- Townsend SW, Manser MB (2012) Functionally referential communication in mammals: The past, present and the future. *Ethology* 118:1–11. doi: 10.1111/eth.12015
- Venables WN, Ripley BD (2002) *Modern Applied Statistics with S, Fourth*. Springer, New York
- Vignal C, Mathevon N, Mottin S (2008) Mate recognition by female zebra finch: Analysis of individuality in male call and first investigations on female decoding process. *Behavioural Processes* 77:191–198. doi: 10.1016/j.beproc.2007.09.003
- Van de Waal E, Bshary R (2010) Social-learning abilities of wild vervet monkeys in a two-step task artificial fruit experiment. *Animal Behaviour*
- Van de Waal E, Claidière N, Whiten A (2013) Social learning and spread of alternative means of opening an artificial fruit in four groups of vervet monkeys. *Animal Behaviour* 85:71–76. doi: 10.1016/j.anbehav.2012.10.008
- Van de Waal E, Renevey N, Favre CM, Bshary R (2010) Selective attention to philopatric models causes directed social learning in wild vervet monkeys. *Proceedings of the Royal Society B: Biological Sciences* 277:2105–2111. doi: 10.1098/rspb.2009.2260
- Ward J (1963) Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* 58:236–244.
- Weihls C, Ligges U, Luebke K, Raabe N (2005) klaR Analyzing German Business Cycles. In: Baier D, Decker R, Schmidt-Thieme L (eds) *Data Analysis and Decision Support*. Springer-Verlag, Berlin, pp 335–343
- Weilgart L, Whitehead H (1997) Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whales. *Behavioral Ecology and Sociobiology* 40:277–285.

- Weisrock DW, Smith SD, Chan LM, et al. (2012) Concatenation and concordance in the reconstruction of mouse lemur phylogeny: an empirical demonstration of the effect of allele sampling in phylogenetics. *Molecular Biology and Evolution* 29:1615–1630.
- Weiss DJ, Newport EL (2006) Mechanisms underlying language acquisition: Benefits from a comparative approach. *Infancy* 9:241–257.
- Welbergen JA, Davies NB (2008) Reed warblers discriminate cuckoos from sparrowhawks with graded alarm signals that attract mates and neighbours. *Animal Behaviour* 76:811–822. doi: 10.1016/j.anbehav.2008.03.020
- Wertheim JO, Worobey M (2007) A challenge to the ancient origin of SIVagm based on African green monkey mitochondrial genomes. *PLoS Pathogens* doi: 10.1371/journal.ppat.0030095
- Wheeler BC (2010) Production and perception of situationally variable alarm calls in wild tufted capuchin monkeys (*Cebus apella nigritus*). *Behavioral Ecology and Sociobiology* 64:989–1000. doi: 10.1007/s00265-010-0914-3
- Wheeler BC, Fischer J (2012) Functionally referential signals: A promising paradigm whose time has passed. *Evolutionary Anthropology* 21:195–205. doi: 10.1002/evan.21319
- Wheeler BC, Hammerschmidt K (2013) Proximate factors underpinning receiver responses to deceptive false alarm calls in wild tufted capuchin monkeys: Is it counterdeception? *American Journal of Primatology* 75:715–725. doi: 10.1002/ajp.22097
- Wheeler BC, Searcy WA, Christiansen MH, et al. (2011) Communication. In: Menzel R, Fischer J (eds) *Animal Thinking: Contemporary Issues in Comparative Cognition*. MIT Press, pp 187–205
- Wich S, Nunn C (2002) Do male “long-distance calls” function in mate defense? A comparative study of long-distance calls in primates. *Behavioral Ecology and Sociobiology* 52:474–484. doi: 10.1007/s00265-002-0541-8
- Wich SA, Schel AM, Vries H de (2008) Geographic variation in Thomas langur (*Presbytis thomasi*) loud calls. *American Journal of Primatology* 70:566–574. doi: 10.1002/ajp.20527
- Wilbrecht L, Nottebohm F (2003) Vocal learning in birds and humans. *Mental Retardation and Developmental Disabilities Research Reviews* 9:
- Wild JM (1993) The avian nucleus retroambigialis: a nucleus for breathing, singing and calling. *Brain Research* 606:319–324.
- Wiley RH (1991) Associations of song properties with habitats for territorial oscine birds of eastern North America. *American Naturalist* 973–993.
- Wilkins MR, Seddon N, Safran RJ (2013) Evolutionary divergence in acoustic signals: causes and consequences. *Trends in Ecology & Evolution* 28:156–166. doi: 10.1016/j.tree.2012.10.002
- Winter P, Handley P, Ploog D, Schott D (1973) Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour* 47:230–239.

- Wray A (1998) Protolanguage as a holistic system for social interaction. *Language & Communication* 18:47–67. doi: 10.1016/S0271-5309(97)00033-5
- Wytenbach RA, May ML, Hoy RR (1996) Categorical perception of sound frequency by crickets. *Science* 273:1542.
- Yamaguchi C, Izumi A, Nakamura K (2010) Time course of vocal modulation during isolation in common marmosets (*Callithrix jacchus*). *American Journal of Primatology* 72:681–688. doi: 10.1002/ajp.20824
- Yin S, McCowan B (2004) Barking in domestic dogs: context specificity and individual identification. *Animal Behaviour* 68:343–355. doi: 10.1016/j.anbehav.2003.07.016
- Zahavi A (1977) The cost of honesty (further remarks on the handicap principle). *Journal of Theoretical Biology* 67:603–605.
- Zahavi A (1975) Mate selection: A selection for a handicap. *Journal of Theoretical Biology* 53:205–214.
- Zimmermann E, Vorobieva E, Wrogemann D, Hafen T (2000) Use of vocal fingerprinting for specific discrimination of gray (*Microcebus murinus*) and rufous mouse lemurs (*Microcebus rufus*). *International Journal of Primatology* 21:837–852.
- Zuberbühler K (2008) Audience effects. *Current Biology* 18:189–190.
- Zuberbühler K (2000a) Referential labelling in Diana monkeys. *Animal Behaviour* 59:917–927.
- Zuberbühler K (2000b) Causal knowledge of predators' behaviour in wild Diana monkeys. *Animal Behaviour* 59:209–220.
- Zuberbühler K (2000c) Causal cognition in a non-human primate: field playback experiments with Diana monkeys. *Cognition* 76:195–207.
- Zuberbühler K (2002) A syntactic rule in forest monkey communication. *Animal Behaviour* 63:293–299.
- Zuberbühler K (2003) Referential signaling in non-human primates: Cognitive precursors and limitations for the evolution of language. *Advances in the Study of Behavior* 33:265–307.
- Zuberbühler K (2001) Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus campbelli*. *Behavioral Ecology and Sociobiology* 50:414–422. doi: 10.1007/s002650100383
- Zuberbühler K, Cheney D, Seyfarth R (1999a) Conceptual semantics in a non-human primate. *Journal of Comparative Psychology* 113:33–42.
- Zuberbühler K, Jenny D, Bshary R (1999b) The predator deterrence function of primate alarm calls. *Ethology* 105:477–490.
- Zuberbühler K, Noë R, Seyfarth RM (1997) Diana monkey long-distance calls: messages for conspecifics and predators. *Animal Behaviour* 53:589–604.

## Acknowledgements

I want first to express my gratitude to Julia Fischer, for taking me into her academic family and for offering support and constructive critique in just the right measures along the somewhat meandering PhD path. In particular I would like to thank Julia for her generous sharing of thoughts and ideas during discussions that inspired and shaped my interest in animal communication. Also for her kindness, her enthusiasm, and for persevering that I should learn some German. Previous supervisors also helped me down this road; in particular, Peter Slater gave inspirational lectures on animal behaviour, and Stuart Semple and Kate Arnold shared their enthusiasm for, and knowledge of, vocal communication in primates. Robert Seyfarth, Dorothy Cheney and Tom Struhsaker generously shared their recordings, observations and insights from their many years with the vervets of Amboseli.

I would like to express my thanks to Peter Kappeler for being my co-supervisor and making the time to evaluate this thesis, and to Julia Ostner, Eckhard Heymann, Dirk Semmann and Ralf Heinrich for agreeing to be on my thesis committee. Research in Germany, Senegal and South Africa was made possible by the financial assistance of the DAAD, the Leakey Foundation, and the Wenner Gren Foundation. I am particularly grateful to DAAD for making the integration into Germany such an enriching experience.

I had the great good luck to carry out my PhD within the Cognitive Ethology Lab, and my thanks go to all my fellow Coggies who made the work place such a friendly and idea-inspiring place to work. Thank you to Matthis, Gisela, Barbara and Brandon for their friendship and discussion in and out of work (plus technical assistance, cake provision and Brandon's numerous insightful proof-reads). Thanks also to Philip for his good company and clustering, Tanja for her patience in explaining the importance of correct nomenclature, Laura for lending me her home, Chris for translations and for always knowing how things should be done, Vanessa for organising great courses, Dietmar for always having an open door and friendly ear, Urs for sharing his green monkey stories and predator models, Matthieu for assistance with monkey videos and French, and Rebecca for being a great desk buddy. I am especially grateful to Kurt for giving much "sound" advice, to Ludwig for his amazing patience and efficiency with long lists of field equipment, and to Mechtilde for helping so cheerfully to steer me through the confusing world of German administration and of course also for the after-lunch chocolate. To the other members of the PhD quartet, Rebecca R, Chris and Julie, thank you for providing support, laughter and melon in the often stressful times of the final write-up. Janna, Katha, Ulli, Cedric, Sally, Marlies, Pascal, Christof, Adeelia and Fede, thank you for making Göttingen a home.

For field seasons in Senegal I would like first of all to thank the Direction de Parcs Nationaux du Senegal for their permission to carry out my research within Niokolo Koba National Park; particular thanks go to Mamadou Sidibe for his help and advice during his time as Conservateur du Parc, and to Chef de Zone Moussa N'Dour. During my stay Simenti Coggies Peter and Annie

were kind field mates who made me welcome at the field site and introduced me to Senegal life, Matze provided very timely help in the field, Janna was excellent company for martial eagle modelling and sharing a gazelle, and Nele gave expert scientific and artistic assistance. My research could not have been carried out without the professional help of a team of Senegalese field assistants; especial thanks go to Jacky, Ciss and Sane for their dedication to the work, which often included hot and frustrating hours waiting for the right moment for field experiments, which very often did not arrive. Heartfelt thanks go to Taye for good company and delicious cuisine. In addition, time spent with the hotel crowd and Simenti people, in particular Boudiang, Mody, Mor, Faude, Boudou, Ferdinand, Fall, Aniset, Diatta, Sagna and Ndao taught me important skills in wolof, belote and the making of tea. Many pleasurable hours were whiled away in Tamba and Dakar in the company of Ace, Hanifatou, Amadou, Beau, Sylla, Absa, Waz, Ba, Saleou and Madiou. Especial thanks go to Colvin and Ndaye Coumba for opening up their home and being such great company, to Reamon for homemade lemonade, and to Ruth and Ibra for providing such a friendly safe haven in Dakar. I thank Samba Kaita, Mansour Mbaye, Ibou, Li, Zack and Claire for providing expert help coupled with good company in the planning and carrying out of project work for waste management and environmental education.

For field work carried out in South Africa, I am grateful to the Mpumalanga Parks Board and ABEERU of UNISA for their permission to work at Loskop Dam Nature Reserve, to Tersia de Beer and Alan Barrett for their help with administration, to Daniel van der Post for sharing his recordings, and to Jannie and Elizma Coetzee for their kindness and hospitality on many occasions. Thanks especially go to the Loskop crew, Nicole, Earl, Chris, Viviane, Matthias, and Oumar, to co-ordinate field work and travel 7 with field gear in a two-seater at the end of a sweaty field day and remain friends is no mean achievement, Danki for the good company, card games and braais.

I want to express especial thanks to my mother, Charlotte, for inspiring me with a love of nature, for believing in whichever life choices I made, and for sending me birthday cake and newspaper cuttings no matter where in the world I was. I want also to thank Will, Topper and Annie for putting up with fleeting visits and still always making time, Tessa for 20 years of beautiful monkey cards, Milly for long letters of home, Stephen for technical advice, and Rachael, Aoife, Caitlin and Freya for their numerous postcards and pictures. Louise, Cathy, Ethan, Chianna, Caroline, Petra, Andy, Monica, Lara, Roxy, Nat and Senta, thank you for putting up with an often distant friend. Finally, thank you Oumar for sharing the many ups and downs with joie de vivre, kindness, patience, good stories and not so good singing. Magui lley gerem di la sant si diegue bi ga ma diégué donté kii ma soriga.

---

---

I hereby declare that I have written this thesis independently and with no other aids or sources than quoted

---

Göttingen,

Tabitha Price



