HAL open science

Animal linguistics: a primer

Mélissa Berthet, Camille Coye, Guillaume Dezecache, Jeremy Kuhn

To cite this version:

Mélissa Berthet, Camille Coye, Guillaume Dezecache, Jeremy Kuhn. Animal linguistics: a primer. Biological Reviews, 2023, 98 (1), pp.81-98. 10.1111/brv.12897 . hal-03875869

HAL Id: hal-03875869

https://hal.science/hal-03875869v1

Submitted on 14 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

BIOLOGICAL REVIEWS
Cambridge Philosophical Society

Biol. Rev. (2022), pp. 000–000.
doi: 10.1111/brv.12897

Animal linguistics: a primer

Mélissa Berthet[1,2,3,†,*], Camille Coye[1,4,†], Guillaume Dezecache[5] and Jeremy Kuhn[1]

[1] *Institut Jean Nicod, Département d'études cognitives, ENS, EHESS, CNRS, PSL University, 75005, Paris, France*
[2] *Center for the Interdisciplinary Study of Language Evolution, University of Zürich, Affolternstrasse 56, 8050, Zurich, Switzerland*
[3] *Department of Comparative Language Science, University of Zürich, Affolternstrasse 56, 8050, Zurich, Switzerland*
[4] *Center for Ecology and Conservation, Bioscience Department, University of Exeter, Penryn Campus, Penryn, TR10 9FE, UK*
[5] *Université Clermont Auvergne, LAPSCO, CNRS, 63000, Clermont-Ferrand, France*

ABSTRACT

The evolution of language has been investigated by several research communities, including biologists and linguists, striving to highlight similar linguistic capacities across species. To date, however, no consensus exists on the linguistic capacities of non-human species. Major controversies remain on the use of linguistic terminology, analysis methods and behavioural data collection. The field of 'animal linguistics' has emerged to overcome these difficulties and attempt to reach uniform methods and terminology. This primer is a tutorial review of 'animal linguistics'. It describes the linguistic concepts of semantics, pragmatics and syntax, and proposes minimal criteria to be fulfilled to claim that a given species displays a particular linguistic capacity. Second, it reviews relevant methods successfully applied to the study of communication in animals and proposes a list of useful references to detect and overcome major pitfalls commonly observed in the collection of animal behaviour data. This primer represents a step towards mutual understanding and fruitful collaborations between linguists and biologists.

Key words: semantics, meaning, syntax, pragmatics, communication, signal.

CONTENTS

I. Introduction . . . 2
II. Definitions and concepts . . . 2
(1) Semantics . . . 3
(*a*) Semantic denotation: the core meaning . . . 5
(*b*) Pragmatics: the contextual meaning . . . 6
(*c*) Communicating information about a signaller's permanent traits . . . 8
(*d*) The case of deception . . . 8
(2) Syntax . . . 9
(*a*) Step 1: detecting syntactic structures . . . 10
(*b*) Step 2: qualifying syntactic structures . . . 10
(*c*) Assessing the semantic value of the component parts . . . 12
(*d*) Hierarchical structures . . . 12
III. Analysing animal linguistic data: a toolbox . . . 12
(1) Establishing a signal repertoire . . . 13
(2) Computational linguistics to detect syntactic patterns . . . 13
(3) Apparently Satisfactory Outcome to investigate semantics . . . 13
(4) Modelling meaning with truth/applicability conditions . . . 14
(5) Principles of competition in pragmatics . . . 14
(6) Collecting animal data: where to start? . . . 15
IV. Conclusions . . . 15
V. Acknowledgements . . . 15

* Author for correspondence (E-mail: melissa.berthet.ac@gmail.com).
† Authors contributed equally to this work.

Biological Reviews (2022) 000–000 © 2022 The Authors. *Biological Reviews* published by John Wiley & Sons Ltd on behalf of Cambridge Philosophical Society. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

VI. References 15
VII. Supporting information 18

I. INTRODUCTION

How language evolved is a long-standing question in science (Christiansen & Kirby, 2003). To answer this question, one fruitful strategy is to break human language down into various component abilities (Hauser, Chomsky & Fitch, 2002; Fitch, 2005). The phylogenetic distribution of each individual component can then be investigated, by comparing communicative capacities across species (Hauser *et al.*, 2002). This leads to the identification of homologies (traits inherited from a common ancestor) or analogies (traits that fulfil a similar function, but which have evolved independently). Species that are phylogenetically close to us (e.g. non-human primates) can therefore be studied to understand the evolutionary history of a human capacity. Studies on phylogenetically more distant species (e.g. birds) can help us to understand the selective pressures that acted on our ancestors and favoured the evolution of human communication as it exists today (Fitch, 2015).

In recent years, significant progress has been made in understanding communicative abilities across a number of species (e.g. Searcy, 2019). However, the interpretation and linguistic relevance of these capacities remains heavily debated (see, for example, Hauser *et al.*, 2002; Scott-Phillips, 2015*b*; Schlenker *et al.*, 2016*b*; Suzuki, Wheatcroft & Griesser, 2018; Bolhuis *et al.*, 2018). In some cases, disagreements originate from fundamental differences in the approach, methods and technical vocabulary used by researchers involved in purely linguistic agendas, and those working on the communication of non-human animals. This is unsurprising: the human linguistic capacity is an easily observable phenomenon into which we have introspective judgments (e.g. whether an utterance is natural, and when it can be used; Bolinger, 1968; Marantz, 2005; Sprouse, 2013) which can be investigated in a relatively direct manner (e.g. we can ask humans about their own practices). Animals, on the other hand, possess their own species-specific perception of the world, cognitive capacities and processes, and communication abilities: these cognitive phenomena are only accessible to human observers *via* measures of behaviour, using ethological methods (Olmstead & Kuhlmeier, 2015). As a result, field-specific terminology has emerged: the same term can be used to describe slightly different concepts in linguistics and biology (for example, the word 'syntax'). Moreover, the great differences between biological and linguistic methodologies often make direct comparisons of results extremely challenging. In one striking example, Prat (2019) argued that while ethological methodologies have thus far had little success in finding 'language' in non-human animals, applying the same methodologies to humans also fails to find any sign of 'language' in human communicative behaviour. The few attempts at direct communication or collaborative efforts between biological and linguistic fields have sometimes been highly technical (and thus of limited accessibility) or dismissively critical, thus discouraging further constructive exchanges. As a result, comparisons between human language and animal communication have often been considered unfruitful and inefficient, and collaboration bound for failure.

Many of these difficulties can be overcome. This can be achieved by increasing exchanges and collaborations between fields, unifying methods and terminology and improving the relevance of comparisons between human and animal communicative systems.

There has been a recent increase in collaborative efforts between linguists and researchers studying animal communication. These have included the search for computational properties of language in other vocal and gestural communication systems (e.g. Heesen *et al.*, 2019), and the application of formal linguistic approaches (e.g. Schlenker *et al.*, 2016*b*,*c*) or computational linguistic approaches (e.g. Kershenbaum *et al.*, 2014*a*; Leroux *et al.*, 2021) to primate vocal communication. These efforts have shown that the use of linguistic concepts in comparative research with animals is both possible and fruitful.

However, these bridges are still fragile, and are only crossed by a handful of researchers. As a way to further encourage this enterprise, we offer here a primer to establish strong basic foundations for animal linguistics. In particular, we aim to provide linguists with the tools to study animal communication, and to provide biologists with basic linguistic notions applicable to the study of animal communication, using concepts and criteria compatible with modern linguistic thinking. This primer is the product of a collaboration between researchers on animal communication and linguists. It can be read as a guide for students and researchers of biology and linguistics alike: first, we define the linguistic concepts of semantics, pragmatics and syntax, in a way that is both biologically and linguistically relevant; second, we present data-analysis methods that have already been successfully applied to animal systems to investigate their linguistic properties. We also include a list of introductory readings for linguists interested in working with animal communication data. A large number of studies in animal linguistics focus on primate vocalizations, which consequently represent an important part of the examples presented here. However, this guide is intended to be applicable to all species and communication modalities; we thus encourage readers to investigate the species of their choice and to consider communication capacities outside the vocal domain.

II. DEFINITIONS AND CONCEPTS

Behaviours of humans and non-human animals can be explained by a variety of cognitive mechanisms, and similar

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

mechanisms can be seen as convincing evidence of evolutionary continuity between humans and other species. However, establishing analogies between human and animal processes involves ruling out alternative explanations. One useful criterion is Morgan's Canon, which states that behaviours should not be interpreted as resulting from higher cognitive faculties (e.g. theory of mind) if they can be interpreted as the outcome of lower capacities (e.g. associative learning) (Shettleworth, 2010). This principle prevents researchers from assigning human-like (and supposedly cognitively more complex) capacities to animals without first rejecting alternative hypotheses ('anthropomorphism'). The field of animal linguistics greatly benefits from the systematic application of Morgan's Canon by researchers, but the lack of unity in definitions, misunderstandings of linguistic concepts, and misuse of linguistic terminology has led to highly debated claims, despite researchers' efforts to avoid anthropomorphism.

In this section, we provide precise definitions of the main linguistic concepts (semantics, pragmatics, and syntax; see Table 1 for a summary of core concepts), using general principles that can be applied equally well to human and non-human communication. For each concept, we provide lists of criteria that can be used to evaluate a species' linguistic capacity.

In the present paper, we restrict our scope to the study of signals. Signals are potential sources of information that are plastically produced at a cost in response to changes in the environment, and are improved over evolutionary time to best fulfil their communicative function (Bradbury & Vehrencamp, 2011). Signals can take different forms: vocalizations, facial or body movements, chemosensory signals, *etc*. Signals contrast with two other sources of information, sometimes called 'signs' and 'cues'. Signs, as defined in the biological literature, are also evolutionarily shaped to convey information, but they do so in a permanent fashion (e.g. warning colours) (Hauser, 1996) (note that this is a distinct concept from the definition of 'sign' used in semiotics and philosophy of language, see Pierce, 1931). Cues are often generated for purposes other than communication: for example, a footprint in the mud conveys information about the presence of a leopard, but it is not evolutionarily optimized for communicative purposes.

Table 1. Summary of definitions and concepts in animal linguistics.

Concept	Definition
Meaning	The set of features of circumstances that appear at a rate greater than chance across the signal's occurrences.
Semantic denotation	The largest set of meaningful features of circumstances that appear across all occurrences of the signal.
Pragmatic inference	The meaningful features of circumstances that always appear when the signal is emitted in the presence (or absence) of a given contextual feature.
Syntax	The set of rules that determine what sequences are well formed.
Compositional syntax	A system in which the meaning of a syntactic structure is derived from the meaning of its parts.
Non-compositional syntax	A system in which the meaning of a syntactic structure is not derived from the meaning of its parts.

(1) Semantics

Semantics pertains to the meaning of a signal. In its largest sense, semantics investigates both the core content of an expression (later referred to as 'semantic denotation'), but also the additional inferences that an expression may have in different contexts ('pragmatics') (Chierchia & McConnell-Ginet, 1996). For example, upon hearing the sentence 'It is raining', one understands that there is water falling from the sky, but one may make further inferences depending on the context of emission: for example, the speaker may signal that the laundry should be brought inside, or that their interlocutor should take an umbrella.

A common distinction is the opposition between signals that are symbolic parts of a code *versus* signals that merely correlate with states of affairs [see Grice (1957) on 'non-natural' and 'natural' meaning]. An intuitively clear example is the difference between the utterance 'I'm happy', and the act of smiling. These two signals appear in similar contexts – namely, when the signaller is happy. Nevertheless, there is an intuitive difference between them: the utterance is a symbolic code that 'stands for' a particular state – it is intentionally uttered to convey a message – while the smile is not. According to some (e.g. Scott-Phillips, 2015*a*), only symbolic utterances can be considered meaningful and deserve to be linguistically investigated. But non-symbolic signals can be interpreted by others: a person may react or reply to a smile in a similar way to the uttered sentence. These signals can also be audience-aware: a person may choose to cover or repress their smile if another individual is present. It is still unclear to what degree animal signals can be considered symbolic. For example, some non-human primates have been shown to adapt their gestural behaviour to the attentional state of the audience (Maille *et al.*, 2012) while the alarm signal of crested pigeons (*Ocyphaps lophotes*) is produced by the physical properties of the feathers; the sound occurs when the pigeons flap faster to escape predators, irrespective of the audience (Murray, Zeil & Magrath, 2017). On the whole, while it may be possible to precisely characterize related concepts like audience-sensitivity, it is not clear that there is any well-defined way to characterize 'symbolic' meaning, especially when one moves to the domain of animal communication. One solution to this problem that has been used with good results is thus to use a relatively broad characterization of meaning (Schlenker *et al.*, 2016*b*) that is not restricted to symbolic signals, and to describe the properties of these signals on a case by case basis.

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In animal communication, the typical pattern of signal emission is presented in Fig. 1. An individual with its specific characteristics witnesses a noteworthy event in its surroundings. This event elicits a temporary emotional or physiological state in this individual. The individual emits a signal, which is perceived by a receiver, who reacts by exhibiting a behaviour. During or after the signal emission, the signaller also displays a behaviour. Both the signaller's and receiver's behaviours induce changes in the original event: the situation at the end of the first signal-loop is not exactly the same as before (e.g. all group members are now further from a predator, or a receptive female is closer to the signaller) – it is now a new event that can trigger a new chain reaction and elicit the emission of a new signal.

The emission of each signal is thus associated with a set of circumstances: the external event, the permanent traits of the signaller, the transient state of the signaller, the receiver's behavioural reaction, and the signaller's behaviour. Each of these circumstances is characterized by a specific set of features, which may vary between signal occurrences: e.g. the type, size, shape or distance of the external object, the sex or identity of the signaller, the valence or arousal of the emotional state of the signaller, the type or strength of the signaller's or recipient's behaviour. For animal communication, we define the meaning of a signal as the set of features of circumstances that appear at a rate greater than chance across the signal's occurrences (adapted from Dezecache & Berthet, 2018). This allows one to assign meaning to a signal based on (*i*) the features of circumstances in which it is used, in comparison to (*ii*) the features of circumstances in which it is not. For example, if there is an overall 2% chance that a leopard is present at any given moment, but when a given signal is uttered this chance goes up to 90%, then the feature 'presence of a leopard' is part of the meaning of the signal.

It must be noted that this definition applies primarily to propositional meaning (meaning similar to that of complete sentences in human language, like 'I am happy'). While this provides a necessary first step into a basic understanding of animal communication, further types of meaning may be necessary to analyse signal combinations, for which it will ultimately become crucial to analyse the meaning of the individual parts of the utterances (e.g. the meaning of 'I', 'am' and 'happy'). To this end, we will return to the meaning of the component parts in Section II.2.*c*, once we have introduced the notions of syntax and compositionality.

Finally, a common mistake is to confuse meaning and function of a signal. The meaning of a signal belongs to the proximate level of explanation, relating to the situations or behaviours that directly trigger a communicative signal. In contrast, the ultimate explanation of a signal relates to the

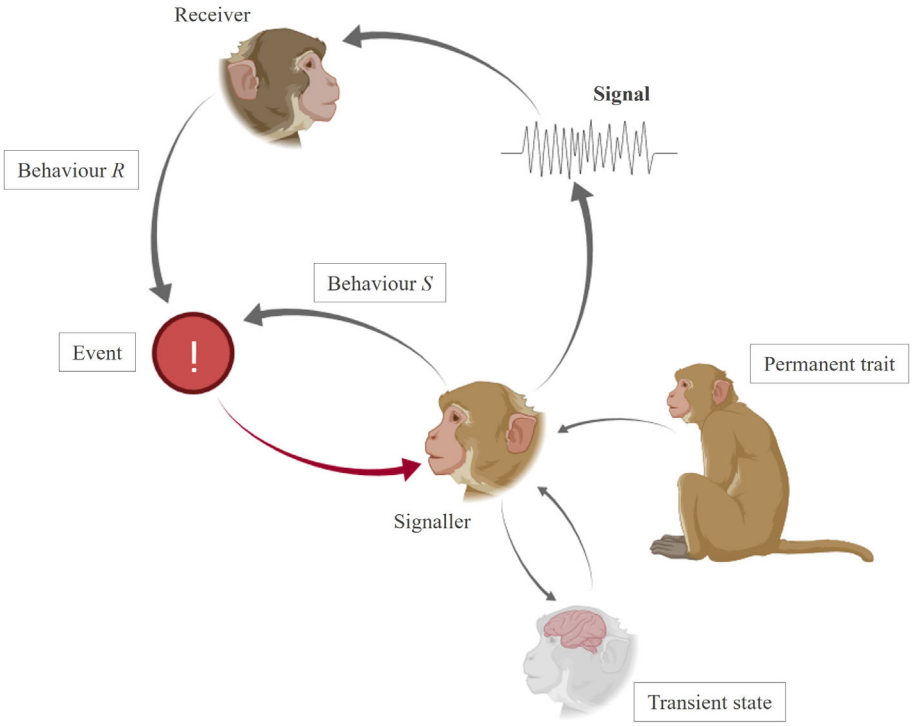

Fig. 1. Semantics summary. A noteworthy external event is witnessed by an individual with specific permanent traits (i.e. long-term characteristics). This event elicits transient states (i.e. temporary emotional and/or physiological reactions), elicits the emission of a signal (here represented by a waveform), and elicits a behavioural response (Behaviour *S*) by the signaller. The receiver performs a behaviour (Behaviour *R*) in response to the signal. The emission of the signal is thus associated with a set of circumstances (represented by the boxes). The semantics of the signal is the set of features of circumstances that appear at a rate greater than chance across the signal's occurrences.

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

adaptive function that it serves on an evolutionary scale. For example, the (proximate) meaning of an alarm signal might be 'there is a predator' but its (ultimate) evolutionary function is to attract attention of the social partners, due to its sharp acoustic parameters. A contact call can mean 'I have pacific intentions' but conveys precise information about the identity and spatial position of the caller to increase inter-individual recognition and facilitate group cohesion. Similarly, some signals, like some birds' songs, do not appear to be meaningful, while their function is to display the readiness of the emitter to defend its territory (Berwick *et al.*, 2011).

(*a*) *Semantic denotation: the core meaning*

In the semantics of human language, the 'denotation' of a word or utterance is commonly defined as its stable semantic contribution (Frege, 1892, 1952; Grice, 1957; Katz & Fodor, 1963). Here, 'stable' indicates that, while a given utterance may be used in a variety of different contexts for a variety of different purposes, the core meaning of the utterance – what it denotes – is that part of meaning that always stays the same, i.e. what is common across all its many uses. For example, although we have seen that 'It's raining' may be used to communicate different things (e.g. that the laundry should be brought inside), the core meaning of the sentence is just that there is rain.

When applied to animal communication, the semantic denotation of a signal is the largest set of meaningful features of circumstances that appear across all occurrences of the signal. For example, if a signal is produced 70% of the time in response to leopards and the other 30% of the time in response to eagles, the denotation of the signal is the set of features common to all occurrences: presence of a predator.

Signals can denote specific features of transient states in the signaller, such as the emotional valence or the type and intensity of physiological states (fear, hunger, sexual receptiveness *etc.*). This seems to be the case for the alarm calls of vervet monkeys (*Chlorocebus pygerythrus*), which are produced in response to specific classes of predators but also during aggression events, probably because different situations elicit similar emotional states in the callers (Price *et al.*, 2015; but see Schamberg, Wittig & Crockford, 2018).

Signals can also denote features of the receiver's behavioural responses (e.g. activity, latency to react, direction or distance of movement, *etc.*). These signals aim at eliciting a specific response in the receivers, i.e. they are goal-directed signals (Schamberg *et al.*, 2018). For example, in a group of wild chimpanzees (*Pan troglodytes*), all usage of the 'present climb on' gesture results in the receiver climbing on the signaller (Hobaiter & Byrne, 2014). Arguably, another example comes from the alarm calls of putty-nosed monkeys (*Cercopithecus nictitans*): females emit 'chirps' to recruit males into predator-deterrence behaviour, and stop calling when the male spots the predator and starts mobbing it (Mehon & Stephan, 2021).

Signals can denote features of the signaller's behaviour (e.g. activity, latency to react, direction or distance of movement, *etc.*), exhibited during or after the emission of the signal. This is the case for signals that are an indication of one's intentions (Cheney & Seyfarth, 2018). In hierarchical social groups, interactions can be ambiguous, and signals can be emitted to reduce uncertainty about the behaviour or intentions of the signaller. In chacma baboons (*Papio ursinus*), females emit grunts when approaching other females, which conveys information about the pacific intentions of females (Silk, Seyfarth & Cheney, 2016). In putty-nosed monkeys, males emit 'pyows' in response to females' alarm calls, while approaching the rest of the group but before spotting the predator: they advertise their engagement to defend the group against a predator (Mehon & Stephan, 2021).

A signal can also denote features of the external event, which comprises the presence of an object (e.g. a predator, specific food, a receptive female), a social interaction (e.g. inter-group aggression), or a situation (e.g. proximity to the territorial border). An example is the alarm 'hoos' of chimpanzees, which denote ambush threats: the presence of snakes elicits the emission of alarm 'hoos', but emission can be suppressed when receivers are already informed, suggesting that the permanent traits and transient states of the caller, the behaviour of the recipient or the behaviour of the caller are not denoted by these calls (Crockford *et al.*, 2012; Crockford, Wittig & Zuberbühler, 2017; see also Girard-Buttoz *et al.*, 2020).

Finally, signals can denote features of the permanent traits of the signaller, i.e. characteristics of the signaller that remain unchanged over long periods of time, such as identity, sex or age class. A convincing case is that of bottlenose dolphins (*Tursiops truncatus*), which use whistles that are individually distinctive and thus, convey individual identity. Importantly, these dolphins can use each other's whistles to address other individuals, and recognize the signature whistle of conspecifics artificially modified to remove voice characteristics (Janik, 2000; Janik, Sayigh & Wells, 2006). We discuss the special issues involved in signalling permanent traits in Section II.1.*c*.

The semantic denotation of signals can vary between species: we do not expect all non-human species to adopt one unique pattern, since different species face different ecological and social pressures and need signals to communicate about a large variety of things. Moreover, we do not expect one species to adopt the same strategy for all the signals of its repertoire: different signals given by a single species may have different semantic denotations because they fulfil different functions (e.g. the social call A of species X can denote features of signaller's intentions while its alarm call B can denote features of external events).

Of course, it is sometimes difficult to disambiguate the semantic denotations of signals as cognitive mechanisms involved in the signalling remain poorly understood, and features of circumstances are often closely correlated. For example, while the emission of signal X may be strongly correlated with the presence of a predator Y, it is difficult to firmly

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

establish that X denotes Y (e.g. 'there is a leopard'), and not the emotional state associated with Y ('I am scared'), the invitation to produce the adaptive response to Y ('run away') or the intention of producing this adaptive response ('I will run away'). In some cases, the precise semantic denotation of a signal can be deduced by careful observations and experimental tests of all the situations and behaviours associated with the emission of the signal, conducted on different populations living in different socio-ecological conditions (e.g. Schlenker *et al.*, 2014). This approach can be complemented with direct (e.g. Liao *et al.*, 2018; Mocha & Burkart, 2021) and indirect (e.g. Schehka & Zimmermann, 2009) investigations of the signaller's transient state. An alternative approach to clarify the exact semantic denotation of a signal is to explore the mental representations it triggers in receivers (if any). This approach has been successfully implemented in primates (Zuberbühler, Cheney & Seyfarth, 1999) and birds (Suzuki, 2018) using experimental protocols. Finally, comparing the semantic denotation of distinct signals in the repertoire can disambiguate the meaning of a given signal (Schlenker *et al.*, 2016*b*; see Section III.5 for more details). However, even after extensive research, it may remain impossible to decide between several hypotheses, because of human or technical limitations.

(*b*) *Pragmatics: the contextual meaning*

In Section II.1, the semantic denotation of a signal was defined as its stable semantic contribution – that part of its meaning that stays the same across all signal occurrences. In contrast, pragmatics pertains to those aspects of meaning that are not stable and depend on context. A child who says 'I need to pee' to their parents while riding in a car is asking them to find a place to stop, while a teenager who utters the same sentence to a sibling taking a long shower is communicating the message 'Get out of the bathroom!'. The semantic denotation of the sentence does not change, but different inferences are made depending on the context in which it is uttered.

For pragmatic inferences, as for the semantic denotation, meaning is defined with respect to the features of a circumstance that appear at a rate greater than chance, comparing situations in which the signal is used to situations in which it is not. However, pragmatic inferences depend on contextual features. Pragmatic inferences are the meaningful features of circumstances that always appear when the signal is emitted in the presence (or absence) of a given contextual feature. Pragmatic inferences are not part of the semantic denotation: they are elicited by variations of contextual features, and enrich the meaning of the signal beyond its semantic denotation (see Fig. 2).

In the example above, the signal 'I need to pee', when uttered in the car, is meaningfully associated with a feature of the receiver's behavioural response – the parent is significantly more likely to stop the car than if the child had said nothing. On the other hand, this behavioural response disappears if the child is talking to their sibling in the shower. The receiver's behaviour 'stop the car' is thus a pragmatic inference that is elicited by the contextual feature 'where is the sentence uttered?', but it is not part of the semantic denotation.

Pragmatic inferences involve reasoning both for the signaller and the receiver. For the receiver, the question is: 'why was this particular signal used and not another, and why was it uttered now?' For the signaller, it is the opposite question: 'which signal should I use, and when?' Because pragmatics involves reasoning about communicative acts, it is often taken to interact with theory of mind – that is, the ability to entertain theories about why others do what they do, for instance by attributing to them some mental state (Premack & Woodruff, 1978). This ability is implicitly present in one way of describing pragmatics: 'what is the speaker trying to say?' However, as highlighted by Schlenker *et al.* (2016*b*), pragmatic reasoning does not have to require a theory of mind. Instead, it can rely on simple associative learning. For example, if an office worker hears people running down the hallway, there are a number of possible explanations: there could be a fire or there could be free pizza. The office worker nevertheless is likely to conclude that there is not a fire, due to the absence of a more specific signal: the fire alarm has not gone off. The office worker is thus reasoning about the state of the world, based on the absence of a signal (fire alarm) which is normally associated with a circumstance's feature (fire); this certainly does not require the reasoner to have a theory of mind of fire alarms. Pragmatic principles relying on strong associations between signals and circumstances' features have been applied to the call systems of several primates, without any assumption about their theory of mind. For example, Schlenker *et al.* (2016*b*) propose that, in male Campbell's monkeys (*Cercopithecus campbelli*), 'krak' calls denote all kinds of predators (aerial or terrestrial), and 'hok' calls denote aerial predators. In many cases, though, recipients infer that a terrestrial predator is present when hearing a 'krak', because if an aerial predator was present, a 'hok' would have been emitted. This reasoning is simply based on the animals' knowledge of their signal repertoire and their circumstances of use, and does not presuppose any theory of mind. This said, accumulating evidence suggests that some non-human animals (e.g. great apes) do possess a theory of mind (e.g. Krupenye *et al.*, 2016; Kano *et al.*, 2019), which they could use for pragmatic reasoning.

Non-human pragmatics should therefore be investigated through one guiding question: what kinds of information can be incorporated and reasoned about in the communication of different species? Some kinds of information may depend on low-level cognitive functions, such as basic perception; other kinds of information may constitute higher level cognitive functions, such as representing the knowledge or intentions of others [see Scott-Phillips (2017) on 'weak pragmatics' *versus* 'strong pragmatics'].

First, pragmatic inferences can be elicited by the presence (or absence) of directly observable contextual parameters. These directly observable contextual parameters can be the external events. For example, in a playback study, female

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

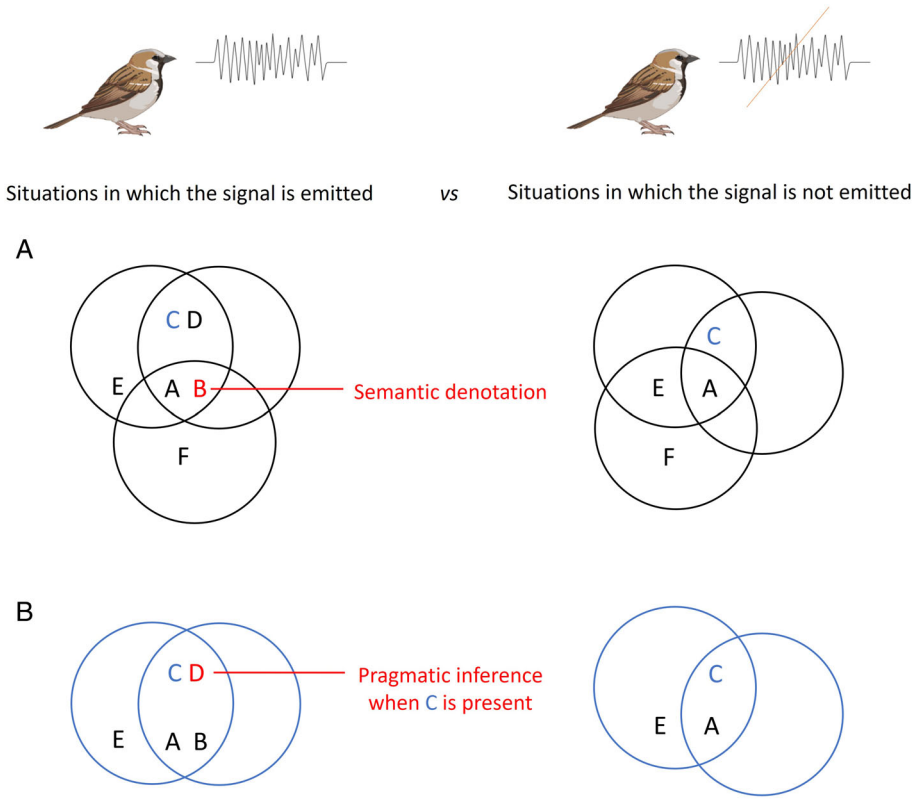

Fig. 2. (A) Schematic representation of the semantic denotation. Each circle represents a situation in which a given signal was emitted (left) or not (right). Each of these situations is characterised by circumstances that have specific features (letters). The semantic denotation of the signal is the largest set of meaningful features of circumstances that appear across all occurrences of the signal: here, only B fulfils this criterion. A is always present, whether or not the signal is emitted: it is not associated with the signal above chance, so it is not a meaningful feature. C, D, E and F are not always present when the signal is emitted, so they are not part of the semantic denotation of the signal. (B) Schematic representation of a pragmatic inference. What pragmatic inferences are elicited by the contextual parameter C? To answer this, we only look at situations in which C is present (blue circles). The inferences of the signal in the context of C are the meaningful features of circumstances that always appear when C is present. Here, only D fulfils this criterion. B is part of the semantic denotation. A is not associated with the signal above chance, so it is not meaningful. E is not always present when the signal is emitted while C is present, so it is not a pragmatic inference of the signal in the context of C. An example of these features in humans, with the signal 'I need to pee': A, the speaker is in good health; B, the speaker's bladder is full; C, the speaker is in a moving car; D, the car stops; E, the speaker is wearing a red T-shirt; F, the speaker's sibling offers a drink to the speaker. An example of these features in putty-nosed monkeys, with the signal 'hack': A, the signaller is in a tree; B, there is a general alert; C, a tree falls; D, the recipient does not look upwards; E, there is a feeding tree nearby; F, the recipient grooms the signaller.

putty-nosed monkeys were shown to react differently to male calls depending on observable properties of the environment, like the presence of noise of a falling tree or acoustic cues of a predator's presence [Arnold & Zuberbühler (2013); see also Arnold & Bar-On (2020) for a discussion]. The semantic denotation of the signal itself remained the same (it is a general alert), but the recipient enriched the meaning of the signal (i.e. it modified its behaviour) based on observable features of the circumstances co-occurring with emission of the signal. Similarly, Diana monkeys (*Cercopithecus diana*) react to conspecifics' alarm calls differently depending on the emission of prior calls or the presence of environmental cues (e.g. previous signs of the presence of a predator) (Zuberbühler *et al.*, 1999). Pragmatic inferences can also be elicited by variations of the signaller's behaviour, like its gaze direction: when an event (e.g. the presence of a predator) elicits the emission of a signal that does not denote the event's location, receivers can retrieve location information from the signaller's behaviour (e.g. infer that the predator is in the canopy if the signaller is looking upwards) (Davidson *et al.*, 2014).

Second, pragmatic inferences can be elicited by the presence (or absence) of contextual parameters that are not directly observable by the signaller or the receiver. These include, for example, the representation of the group's social structure and the memory of past social interactions (e.g. Bergman & Sheehan, 2013; Wittig *et al.*, 2014). For example, chimpanzees react differently to aggressive barks of conspecifics that are closely bonded to a subject's former

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

opponent: individuals can thus enrich a signal's semantic denotation (e.g. an aggressive interaction) with social knowledge (e.g. friendship and social structure of the community) and past personal history (e.g. having been subject to aggression from a conspecific) (Wittig *et al.*, 2014). Another example comes from the representation of third-party relationships: in chimpanzees, victims of severe attacks produce screams whose acoustic structure exaggerates the level of aggression experienced if the audience includes at least one listener whose rank matches or surpasses that of the aggressor (Slocombe & Zuberbuhler, 2007). Finally, these can include representations of the knowledge or belief states of others: for example, wild chimpanzees modulate alarm calling and other communicative behaviour as a function of conspecifics' knowledge (Crockford *et al.*, 2012).

We note that it may sometimes be difficult to disentangle pragmatic inferences from semantic denotations, especially when it is difficult to decide whether two different forms are occurrences of the same signal (see for example Kuhn *et al.*, 2018). For example, black-fronted titi monkeys (*Callicebus nigrifrons*) possess two acoustic variants of the alarm B-call: one higher-pitched variant is given in response to terrestrial predators, and a lower-pitched one when the caller is descending to the ground (Berthet *et al.*, 2018). One hypothesis is that B-calls are two different calls, with different semantic denotations: the lower call means 'I am going to the ground' and the higher call means 'there is a terrestrial predator'. An alternative hypothesis is that B-calls have one semantic denotation, regardless of their acoustic structure (e.g. 'I am afraid'), but that contextual parameters act on their acoustic structure and slightly modify their meaning (e.g. 'I am a little afraid', when the caller is going near the ground, *versus* 'I am very afraid', when a terrestrial predator is present). This question can possibly be answered by investigating whether listeners consider the two variants as graded variations of one signal, or as two different signals (see Section III.1).

(*c*) *Communicating information about a signaller's permanent traits*

It is common for a signaller's characteristics (e.g. identity, size, sex) to influence the shape of the signal (e.g. the body size of the caller influences the fundamental frequency of its calls) and transmit reliable information about the signaller to receivers (e.g. Ey, Pfefferle & Fischer, 2007; Bowling *et al.*, 2017). It may be attractive to consider this information as part of the semantic denotation of the signal, similarly to individual names or age labels in human language. The present approach gives a slightly different perspective on such cases.

First, on the approach here, the semantic denotation of a signal is defined as the largest set of meaningful features of circumstances that appear across all occurrences of the signal. Notably, though, permanent features of the signaller are always present, even when the signal is not emitted; they thus do not appear at a rate greater than chance, so cannot be part of the semantic denotation of a signal. On the other hand, signals with a tautological denotation (i.e. signals that are always true) may have the function of conveying the identity and location of an individual: here, the choice to use the signal may itself generate pragmatic inferences. In English, for example, the sentence 'I'm here' is true no matter who or where the speaker is; based on the present framework the sentence has only a trivial semantic denotation. A speaker may nevertheless decide to use the utterance (instead of remaining silent) to elicit a pragmatic inference in the receiver (e.g. the receiver approaches the voice source). Similar reasoning may apply to contact or territorial songs of animals, which are likely not meaningful but have an attractive or defensive function. This is, for instance, likely to be the case for giant otters (*Pteronura brasiliensis*) whose contact calls function to maintain socio-spatial cohesion and reliably convey caller identity (Mumm, Urrutia & Knörnschild, 2014). Receivers can detect the caller's location and identity, from which they draw pragmatic inferences about appropriate behaviour (e.g. to approach the caller or not).

Second, semantic investigations are conducted across all occurrences of a signal. This implies that the object of study is an idealized, stable shape of the signal that is not impacted by the conditions of production: the semantic denotation of the signal remains the same across all occurrences, regardless of the signaller's traits. A point of contrast can thus be drawn between identity information drawn indirectly from a signal (e.g. David Attenborough saying 'It's me' with his distinctively recognizable voice) *versus* identity information that is part of the denotation itself (e.g. David Attenborough saying, 'It's David Attenborough'). If identity information is part of the denotation itself, the meaning should remain even when the signal is emitted by another individual (e.g. anyone else saying 'It's David Attenborough' to refer to the famous biologist).

As a consequence, few animal systems qualify so far as semantic denotations of permanent traits. Earlier, we mentioned the case of bottlenose dolphin whistles: since these signals convey stable information about permanent traits even when the vocal characteristics of the signaller have been removed (Janik *et al.*, 2006), they can be said to semantically denote identity.

(*d*) *The case of deception*

In the definitions above, we have assumed that all animal signals are produced truthfully. This simplification is made out of necessity: since direct introspective methods are not possible for non-human animals, meaning must be defined (at a first pass) *via* features of the real world. In reality, though, an emitted signal may be false (i.e. the signal is not emitted in the set of circumstances with which is normally correlated), either accidentally or as an attempt to deceive.

Deception occurs when an individual produces a signal whose reception will benefit it at the expense of the receiver. Deception has been observed in a wide range of species. For example, fork-tailed drongos (*Dicrurus adsimilis*) produce alarm calls to threats that can be understood by sympatric species, but they also use the same alarm calls in non-threat

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

contexts to scare away these animals and steal their food (Flower, Gribble & Ridley, 2014). Mantis shrimp (*Gonodactylus bredini*) produce meral spread threat displays to drive off conspecific opponents, even when they are newly moulted and thus vulnerable to attack (Adams & Caldwell, 1990). Tufted capuchins (*Sapajus apella*) sometimes produce alarm calls during feeding events, which elicits an escape reaction from conspecifics and allows the caller to access food (Kean *et al.*, 2017). Notably, such examples (as well as deception in human language) provide a challenge for a framework like the one we have presented above, in which the meaning of a signal is defined relative to the observed circumstances of its use, since in cases of deception these circumstances might not be found. However, we believe that these special cases are not a major limitation.

First, for deceptive communication to be effective, a signal can only rarely be emitted in the wrong context. High rates of unreliable signalling may put selective pressure on receivers to learn that the sender is not trustworthy and ignore their signals (Wheeler & Hammerschmidt, 2013). It is thus likely that, if a species displays deceptive communication, these cases will nonetheless remain rare, with relatively little impact on the statistical evaluation of meaning.

Second, deceptive communication can still provide insight into the meaning of a signal, by using the relationship between the semantic denotation and the pragmatic inferences. Namely, when one lies, one nevertheless expects the recipient to behave as though one were telling the truth (e.g. uttering 'I need to pee', in order to escape a boring class). On the present framework, the receiver's behaviour is generally a pragmatic inference (provided the signal does not denote the receiver's behavioural response). Because a receiver has no way of knowing whether a signal is truthful or not, this particular inference – the receiver behavioural response – will remain constant across signal emissions, even if the semantic denotation is not true. The stability of certain pragmatic inferences in a given context thus provides an avenue to hypothesize about the meaning of a signal. One sees that recipients react as though X were the case, even if it actually is not (this logic underlies the playback methodology, in which one observes reactions of an animal to a false signal). But notably, there is no silver bullet for working backwards from the pragmatics to the denotation: this requires theories of pragmatics (see Section III.5) and theories of animal behaviour, and likely varies from species to species.

(2) Syntax

Syntax describes the set of rules that determine what sequences are well-formed, and what sequences are not. It is a combinatorial system: that is, it combines and orders units into sequences.

In its most general sense, syntax does not require a semantics; that is, neither the units being combined nor the resulting sequence necessarily have to be meaningful. Consider sequences of parentheses that must be first opened and later closed: the sequences ()() and (()) are well formed, but the sequences (() and))((are not. In animal communication, the presence of syntax without a semantic interpretation has been suggested for birdsong: it is possible to describe a set of rules (i.e. a syntax) that describes which sequences of notes are well formed and which are not (Berwick *et al.*, 2011), but neither the individual notes nor the resulting sequences bear distinct meanings on the definitions above.

However, many syntactic systems do interface with semantics. For humans, it has been observed that natural language involves two distinct combinatorial systems, acting at two different levels (Marler, 1977; Pullum & Zwicky, 1988; Collier *et al.*, 2014). At a first level, phonology combines articulatory units, the phonemes (sounds in spoken language or body movements in sign languages) into words. For instance, English phonology determines that 'plimp' is a well-formed sequence (even if it is not a real word), but 'lpipm' is not. At a second level, (sentential) syntax combines words into sentences. The rules of English syntax, for example, determine that 'The bird is singing' is a well-formed sequence, but that 'Singing the is bird' is not. (These levels can be further refined to additionally include morphology, which combines roots and affixes into words, such as sing+ing.) Notably, for human language, sentential syntax interfaces with semantics: words bear meaning, and the meaning of a sentence is derived from the way that these meanings are combined *via* the syntax (e.g. 'Alex ate the chicken' and 'The chicken ate Alex' involve the same units, but receive rather different interpretations).

The term 'syntax' is ambiguously used in the literature: it sometimes refers to combinatorial systems in general (including birdsong and human phonology), or to the specific system that combines words into sentences in human language – to disambiguate, we will call the latter 'sentential syntax'.

The fact that clear formal properties distinguish these two levels of combination in human language allows further distinctions to be made. For example, idioms are expressions that are built *via* the sentential syntax, but whose meaning is not derived from the meaning of their parts. For example, the idiomatic meaning of 'spill the beans' is unrelated to the meaning of 'beans', but it is nevertheless an output of the syntactic system and not the phonological system (i.e. it is not a single word 'spilthabeens'), due to its interaction with the sentential syntax in other ways (such as tense: 'spilled the beans'). Some specific cases complexify the classification of combinations. For example, some sequences that are generated by the phonological system may nevertheless seem to contain the units of sentential syntax or idioms: for example, the word 'candid' can be decomposed into the sounds 'can' and 'did', which are themselves both words, but this is completely accidental. Another difficult case is due to language evolution. For example, the historical etymology of 'daisy' is the idiomatic 'day's eye', derived from sentential syntax, but this has been reanalysed as phonological structure over time.

These examples illustrate that the distinction between phonology, sentential syntax and idioms in human language relies on a clear delineation of their properties, which are well

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

known and understood. For animal communication, such a distinction may be premature, for our lack of understanding of these systems prevents us from drawing a clear delineation between these concepts. As a result, the distinction that is commonly and productively made for animal communication is whether or not a meaningful combination is semantically compositional: that is, whether or not the meaning of the whole is derived from the meaning of the parts (Frege, 1892, 1952). Both compositional and non-compositional combinations have functional and ecological value. Semantically non-compositional combination allows a large and arbitrary vocabulary to be generated by a small set of units (Fitch, 2019). Semantically compositional combination allows simple meanings to combine to produce more complex concepts without needing to memorize each concept individually (Collier *et al.*, 2014). We might thus expect to observe both kinds of combination in animal communication systems.

We propose a two-step procedure to investigate syntactic properties of non-human systems, summarized in Fig. 3. First, we propose two criteria to detect syntactic structures. Second, we present methods to investigate the interface between the syntactic structure and semantics, to qualify the degree of compositionality of the syntactic structure.

(*a*) *Step 1: detecting syntactic structures*

Syntax describes the combination of units into sequences. The first step of the detection of a syntactic structure consists of identifying individual units and the sequences in which they can appear, and verifying that signals that appear in two different sequences are perceived as the same unit (stage 1 in Fig. 3). In other words, is X_1 in the sequence X_1A the same unit as X_2 in the sequence X_2B? One frequent methodology is to test animals' reactions to artificially created sequences, generated by replacing signals from one sequence with signals from a different sequence: if X_1 and X_2 are perceived as the same units, the recipients should react similarly to X_1A and X_2A, and to X_2B and X_1B. For example, male Campbell's monkeys produce 'krak' alarm calls as well as 'krakoo' calls, which appear to be the combination of the 'krak' call with an '-oo' ending (Ouattara, Lemasson & Zuberbühler, 2009*b*). To test whether the 'krak' in both cases is perceived as the same unit, Coye *et al.* (2015) generated artificial 'krak' and 'krakoo' calls by either adding an '-oo' to a 'krak' call or by removing an '-oo' from a 'krakoo' call. The authors found that Diana monkeys, which associate with Campbell's monkeys, responded similarly to both the natural and the artificial calls, showing that the two 'krak's are perceptually the same. Methods for establishing a signal repertoire are discussed further in Section III.1.

Once combinatorial units are identified, the next stage is to investigate the rules of combination (stage 2 in Fig. 3). A typical methodology consists of drawing hypotheses about the rules of combination from a large combination data set then testing their validity by creating artificial combinations that disrupt the hypothesized combination rules: if recipients react differently to disrupted combinations and original combinations, then the rules of combinations matter. One common combination rule is order. In human language, for example, 'the man' is a well-formed noun phrase while 'man the' is not; 'blue sky' and 'sky blue' are both well formed but do not mean the same thing. Many animal communication systems show a similar importance of order in the syntactic system. The Japanese great tit (*Parus minor*) combines alert calls and recruitment calls into an alert–recruitment sequence structured according to an ordering rule: if the alert and recruitment calls are reversed (i.e. a recruitment–alert sequence), then receivers do not react (Suzuki, Wheatcroft & Griesser, 2016). On the other hand, in other animal communication systems, order may be less important (Engesser & Townsend, 2019). For example, alarm sequences of titi monkeys are structured according to rules of proportions of consecutive call types (Berthet *et al.*, 2019*b*), while those of black-capped chickadees (*Poecile atricapilla*) rely on repetitions of elements (Templeton, Greene & Davis, 2005).

(*b*) *Step 2: qualifying syntactic structures*

After establishing cases of syntactic combination, one can determine whether a combination is semantically compositional. As we have seen, human phonology is non-compositional: the meaning of 'candid' is not related to the meanings of 'can' and 'did'. On the other hand, sentential syntax generally is compositional: the meaning of the sentence 'John left' is derived from the meaning of 'John' and that of 'left'.

Determining whether a combination is semantically compositional can only be done on combinations that are meaningful, i.e. that possess a semantic denotation as defined in Section II.1.*a* (stage 3 in Fig. 3). One of the strongest tools to identify compositionality is productivity (Baayen, 1992; Szabó, 2020). Semantic compositionality implies that a signal can be used in different syntactic combinations (e.g. the word 'elephant' can be used in the sequences 'a big elephant' and 'the elephant'), and productivity is what allows one to interpret the signal in all of these combinations, including entirely novel sequences (e.g. 'The one-eyed elephant is eating blue popcorn'). Productivity thus implies that (a) a signal contributes the same meaning in different sequences, and (b) one can produce and interpret novel sequences. For example, the English suffix '-proof' can be combined with many different nouns, always contributing the same meaning ('bullet-proof', 'water-proof', *etc.*), and can even be applied to words that only recently appeared in the English language ('Covid-proof').

Productivity can be identified and quantified in non-human systems, in a two-step procedure. The first step (4a in Fig. 3) is to verify that a signal can be used with the same meaning across different combinations, and identify the possible combinations. This can be achieved through naturalistic observations: the greater the number of combinations in which a signal appears, the greater the productivity

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

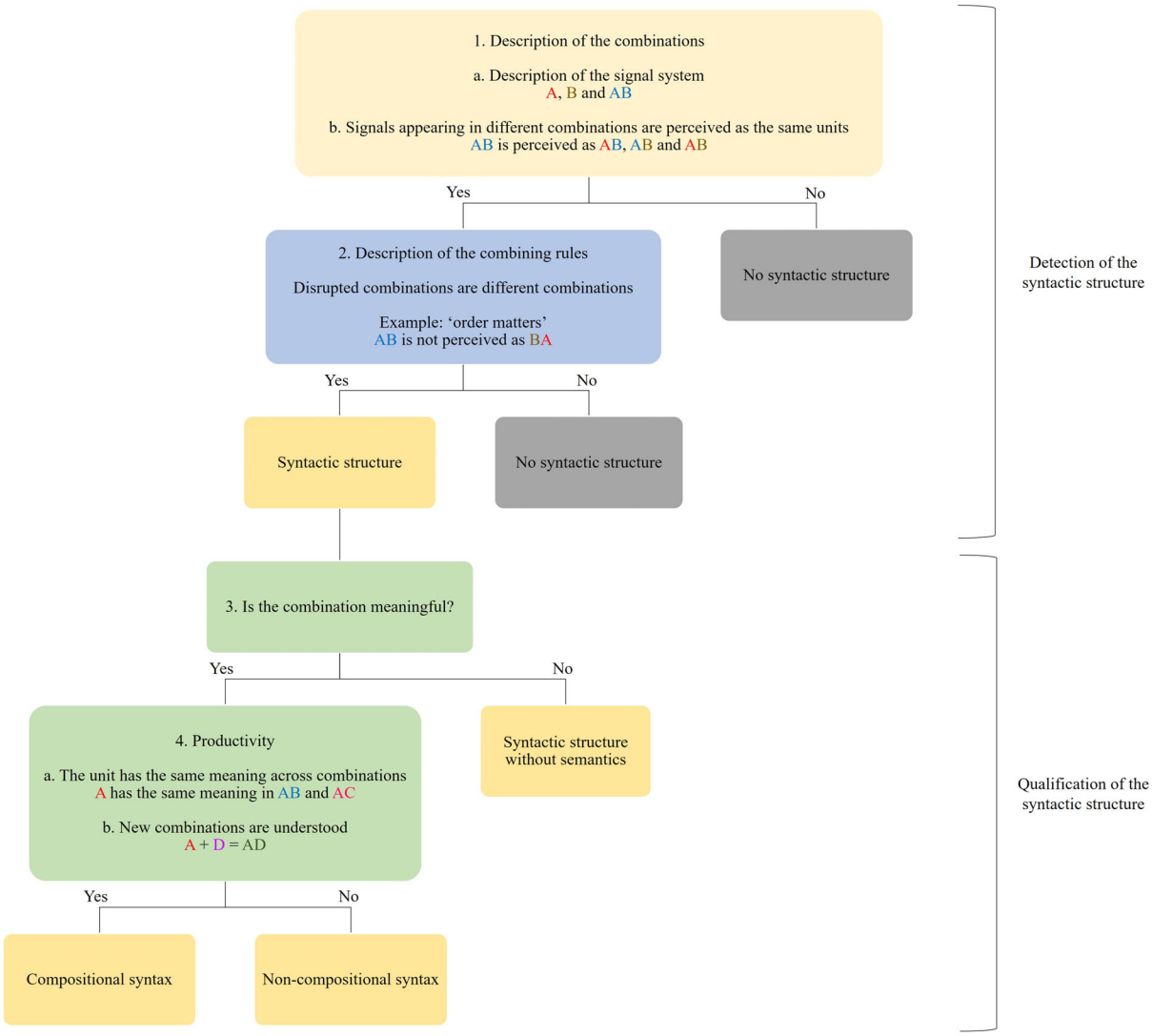

Fig. 3. Detection and qualification of syntactic structures. The analysis steps are illustrated with a fictive system in which two signals A and B can be combined into an AB sequence.

of the system. Campbell's monkeys have two different alarm calls, 'krak' and 'hok', where 'hok' is specific to aerial disturbances (Ouattara, Lemasson & Zuberbühler, 2009*a*). These calls can also be followed by a suffix, '-oo': 'krak-oo' indicates a weak disturbance; 'hok-oo' indicates a weak aerial disturbance (Ouattara *et al.*, 2009*a*). The '-oo' suffix is productive because it contributes the same meaning in two different sequences: in both cases, it attenuates the level of danger. However, the small size of the inventories involved results in relatively weak productivity; an alternative explanation could be that the animals have memorized distinct meanings for each call sequence (Kuhn *et al.*, 2018).

The second step (4b in Fig. 3) is to verify whether non-human animals can understand novel sequences using compositional syntax: if so, the degree of productivity is high. This can be achieved with an experimental paradigm. For example, Japanese great tits produce ABC-calls in response to predators, and listeners respond to this call by scanning the area. D-calls are recruitment calls, emitted in non-dangerous situations to attract the receiver. These calls can be combined into a ABC-D sequence that combines the two meanings: it is emitted in presence of predators to recruit conspecifics for mobbing, and receivers approach the signaller while scanning the area (Suzuki, Wheatcroft & Griesser, 2020). These patterns display at least a weak degree of productivity, since the semantic contribution of both ABC-calls and D-calls is the same across different sequences in which they appear. To investigate a higher degree of productivity, Suzuki, Wheatcroft & Griesser (2017) further artificially combined Japanese tits' ABC-calls with the recruitment call ('tää') of willow tits (*Poecile montanus*), which is used to attract both conspecifics and heterospecifics, including Japanese tits. Japanese tits responded similarly to these entirely novel sequences and to natural ABC-D calls

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

(approach while scanning), thus displaying productivity with novel examples. This high degree of productivity is strong evidence for semantic compositionality.

In contrast, non-compositionality can be detected when the sequence is meaningful but it is impossible to assign a semantic value to the component elements in a way that derives the meaning of the complex sequence. For example, chestnut-crowned babblers (*Pomatostomus ruficeps*) possess two calls composed of the same notes A and B: the flight call (AB) is given during flight and the prompt call (BAB) is given when provisioning nestlings with food (Engesser *et al.*, 2015). Given the great difference in the contexts of use (flight versus feeding nestlings), there is no clear way to derive these two denotations from primitive meanings of the A and B calls.

Another example is the case of the 'pyow-hack' sequences of putty-nosed monkeys (Arnold & Zuberbühler, 2012). 'Pyow' calls are used for general disturbances and attract the attention of receivers, while 'hack' calls indicate eagle presence and inhibit movement. These calls can be combined into a 'pyow-hack' sequence, which elicit group movement in receivers in the absence of predators: the meaning of this combination does not seem to be derived from the meaning of its constituents. Moreover, playback experiments showed that listeners responded similarly to sequences of varying proportion of 'pyow' and 'hack' calls, further suggesting that these combinations are non-compositional [but see Schlenker *et al.* (2016*a*) for another interpretation, discussed further in Section III.5].

(*c*) *Assessing the semantic value of the component parts*

As mentioned in the introduction to Section II.1, our semantic methodology applies primarily to propositional meanings: these correspond to features of circumstances that can be directly observed, and describe facts that can be evaluated as true or false (e.g. 'There is a leopard', 'The signaller is afraid'). In human language, however, there are meaningful signals that do not have this type of meaning – this is the case for most words in isolation. For example, one cannot evaluate the word 'afraid' as true or false without knowing which individual is being described. The word 'not' in isolation also cannot be evaluated as true or false. A similar situation may hold for the compositional syntax of animal communication: even if the meaning of the whole is associated with features of circumstances, the meanings of the parts do not necessarily have this type of meaning (e.g. a signal in a combination could potentially denote negation while the combination as a whole denotes the absence of a predator).

The strategy traditionally adopted by linguists is to establish the meaning of a sentence, and then to work backwards to the meaning of the parts. For example, if we know what 'It's raining' and 'It's not raining' mean, then we can deduce the meaning of 'not'. But even for human language, this methodology leaves room for different theoretical possibilities. For example, in Italian, the sentence 'Nessuno ha visto niente' (literally, 'Nobody has seen nothing') negates the proposition that someone saw something, but it is not immediately obvious which word in the sentence introduces the negative meaning: 'nessuno', 'niente', or something else (Giannakidou & Zeijlstra, 2017). In animals, a similar methodology can be adopted, but the same questions need to be asked, and while our framework allows the debate to be opened, there is currently no one-size-fits-all algorithm to answer them.

(*d*) *Hierarchical structures*

Hierarchical structures are created by syntactic systems in which the output of one combinatorial rule is the input for a second combinatorial rule. In English, for example, 'the' combines with 'bird' to give 'the bird', which can then combine with 'sings' to give the sentence 'the bird sings'. This sentence can be represented hierarchically as [[the bird] sings]. While hierarchy allows the possibility of infinite recursion (e.g. ex-husband, ex-ex-husband, ex-ex-ex-husband, *etc.*), infinite recursion is not a necessary component of hierarchical structure.

Several kinds of evidence have been used to argue for structural hierarchy in the communication of humans and non-human animals. One criterion for hierarchy is the presence of dependencies between non-adjacent elements. In English, for example, the expression 'either … or …' shows a long-distance dependency: the word 'either' must be followed by the word 'or', but the distance between the two can be arbitrarily large (e.g. 'Either you tell me what you told your brother last night or I'll scream'). This dependency can nevertheless be stated by a simple hierarchical rule ('Either S1 or S2' is well formed) that refers to large chunks of structure. Long-distance dependency can be found in canaries (*Serinus canarius*): a syllable type can influence the choice of another syllable type produced up to five syllables later. For example, if a phrase C precedes the phrase sequence DABN, the following phrase is likely to be Y (i.e. sequence CDABNY), while if the phrase N precedes DABN, the following phrase is more likely to be E (i.e. NDABNE) (Markowitz *et al.*, 2013).

Hierarchical structure can also be motivated by identifying constituents, i.e. substrings of a sequence that function as a single unit. For example, in the English sentence 'The bird sang', the substring 'the bird' is a constituent: it behaves as a single unit under a variety of manipulations, including permutations of the sequence ('What sang was the bird') and replacement by other elements ('It sang'). In chimpanzees, the combination of 'hoo' and 'panted hoo' (bigram HO_PH) can be emitted alone but is also found in the larger combinations HO_PH_PS or HO_PH_PB (Girard-Buttoz *et al.*, 2022), suggesting that the bigram HO_PH is a constituent, which can then combine in a larger syntactic frame.

III. ANALYSING ANIMAL LINGUISTIC DATA: A TOOLBOX

Animal communication has long been studied with biological tools, while human language has always been studied with

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

linguistic methods. This lack of unity in methodology makes the two systems difficult to compare. One study recently evaluated human languages with the tools commonly used in animal communication. It failed to highlight semantics, syntactic structure or vocal learning in human language (Prat, 2019). This result is puzzling, and strongly suggests that, in order to properly compare animal and human communication and find linguistic-like capacities in animals, we should unify the methods. We describe below a set of methods that have been successfully applied to both human and non-human communication systems to study semantics, syntax and pragmatics.

(1) Establishing a signal repertoire

The first step of any animal linguistic investigation is to establish a comprehensive repertoire of signals of the species of interest. This step involves observations, measurements, description of the signals and classification. The relevant methodologies have been extensively covered in the ethological literature (see online Supporting Information, Table S1). In humans, one method to verify that two similar signals are the same unit is the contrastive distribution. The typical paradigm consists of replacing one signal by the other, in the same environment (e.g. in the same word). If the signals are not the same unit ('contrastive'), this permutation results in a change in meaning. For example, English has a contrast between /r/ and /l/, which can be highlighted by the fact that the words 'row' and 'low' have different meanings. In Japanese, however, this contrast does not exist: [r] and [l] are variants of a single liquid consonant, so no contrastive pairs can be found.

In animals, contrastive distribution can be used to establish solid signal repertoires based on how animals themselves use and perceive signals: if two signals are contrastive, they elicit a different reaction by the recipient when presented in the same environment.

Contrastive distribution can be used in experimental design to verify that units composing a combination are perceived as the same unit. Chestnut-crowned babblers seemingly combine the A and B notes into the flight combination (A–B structure) and into the prompt combination (B–A–B structure), which have different meanings (Engesser *et al.*, 2015). The authors exposed subjects to natural flight and prompt combinations, and artificially rebuilt combinations (flight combinations made of prompt combination units and *vice versa*). They showed that subjects reacted similarly to natural and artificial combinations, suggesting that this species combines the same units into different combinations (see more details in Engesser *et al.*, 2015). However, this methodology does not necessarily require an experimental design. For example, Hobaiter & Byrne (2017) used an observational paradigm to establish the gestural repertoire of chimpanzees. They showed that wild chimpanzees may swing the arm or swing the leg, but that these two gestures are not contrastive: the two appear in the same circumstances of use so have exactly the same meaning, suggesting that they can be considered as the same unit. In contrast, hitting with the hand *versus* hitting with the foot have different meanings to the chimpanzees, suggesting that these two signals are distinct units.

On the other hand, even when contrastive distributions have been established, there may remain analytical choices for the theoretician. One such example relates to non-compositional syntax, where a syntactically complex signal is considered a single unit at the level of semantics. By our definition of semantic denotation, the English word 'cat' has to mean the same thing each time it is used. But the 'cat' in 'caterpillar' does not have this meaning. This nevertheless does not mean that we should revise our definition of 'cat'; rather, we should revise what is counted as an occurrence of the signal 'cat' to include only cases where it is not followed by '…erpillar'. An exactly analogous situation holds for the non-compositional syntax of chestnut-crowned babblers. The signal AB is used as a flight call, but notably, the same string is a subpart of the BAB prompt call. To establish the meaning of AB as flight-related, one must make the analytical decision to not count occurrences of BAB as instances of the signal AB.

(2) Computational linguistics to detect syntactic patterns

Often, when data sets are large or when patterns are complex, one can experience difficulties identifying combinatorial patterns from observations of the data alone. Several tools derived from computational linguistics methods can help detection and testing for specific patterns in animal communication, such as repetitions, combinations, ordering, overlapping or temporal structures. These computational tools are extensively presented, together with the type of sequences and patterns they are best suited for, in Kershenbaum *et al.* (2014*a*). Of note, Markov models (Kershenbaum *et al.*, 2014*b*; Alger, Larget & Riters, 2016; but see Kershenbaum & Garland, 2015), N-grams models (e.g. Berthet *et al.*, 2019*b*), transitions probabilities (e.g. Jin & Kozhevnikov, 2011) and collocation analyses (e.g. Leroux *et al.*, 2021) can all help highlight which signals are more likely to be combined. String edit distance methods (e.g. Kershenbaum *et al.*, 2012; Kershenbaum & Garland, 2015) can allow long sequences of signals to be compared to detect underlying structure. Hierarchical structure can be investigated using a number of different computational tools, like entropy estimators (Suzuki, Buck & Tyack, 2006), network analyses (Allen *et al.*, 2019), Markovian processes (Sainburg *et al.*, 2019) and quantification of clustering events (Kello *et al.*, 2017), for example.

(3) Apparently Satisfactory Outcome to investigate semantics

Investigating the semantics of signals is a difficult task, especially for signals that occur in a large diversity of contexts. The Apparently Satisfactory Outcome (ASO) is an efficient

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

method to investigate the meaning of signals that are emitted intentionally (Hobaiter & Byrne, 2014), in particular the semantic denotation of goal-directed signals (see Section II.1.*a*), or pragmatic inferences involving receiver's behaviour (see Section II.1.*b*). The ASO is defined as the action performed by the recipient that results in cessation of signalling by the signaller. This method relies on the assumption that, in intentional communication, an individual will continue to emit a signal until the recipient's reaction is congruent with the signal's meaning, i.e. until the reaction is satisfactory to the signaller. This ASO is taken to be the meaning of the signal as intended by the signaller: the semantic denotation and pragmatic inferences of the signal in a population can be derived from ASOs collected across many instances and individuals. This method has been successfully applied to the gestures of apes (including humans) (Graham *et al.*, 2018; Kersken *et al.*, 2018).

(4) Modelling meaning with truth/applicability conditions

In humans, semantics is often investigated using the truth conditions of sentences. Any native speaker of a language knows both whether a sentence sounds natural (syntax) and also what the world must be like for the sentence to be true (semantics). For example, 'The cat sleeps' is true when there is exactly one relevant cat and that cat has the property of sleeping. From sentential truth conditions, a linguist can work backwards to understand the meaning of individual words, by isolating their stable contribution in different sentences. Writing out explicit statements of truth conditions like the following allows sentential truth conditions to be explicitly stated and compared: the sentence S is true exactly when conditions C hold.

It is obviously not possible to ask animals what the meaning of a signal of theirs is. Also, it remains unknown whether any cognitive processes exist in non-human animals that correspond to the notion of 'truth' for sentences of human language. One solution proposed by Schlenker *et al.* (2016*b*) is to investigate the conditions under which signals are applicable and inapplicable using natural observations and experiments. The meaning of a signal can thus be written as: the signal S is applicable exactly when conditions C hold.

This framework allows researchers to draw precise theories about the use and structure of animal's signals, regardless of the cognitive capacities of the species, and derive testable hypotheses about the semantics and syntax of signals. This method can be applied to any type of animal semantics (see Section II.1), including signals whose denotation seems general or unclear (see Dezecache & Berthet, 2018), and has been applied to the vocal systems of several primates (Schlenker *et al.*, 2016*b*; Berthet *et al.*, 2019*a*).

(5) Principles of competition in pragmatics

As discussed in Section II.1.*b*, meaning in human and animal communication is often enriched by pragmatic processes. One important insight in pragmatics is the principle of competition between alternatives: inferences are frequently made based on what could have been said, but was not (Grice, 1957). For example, imagine that Mary and John have a dog named Max. John, looking out the window, says 'A dog is playing in the garden'. When hearing John's sentence, Mary will probably infer that John is not watching Max, but another dog that he does not know. In particular, if the dog in the garden were Max, the sentence 'A dog is playing in the garden' would still be true, but John would be unlikely to say it, because there is a simpler and more informative alternative that he could say instead: 'Max is playing in the garden'. Since he did not say this sentence, Mary infers that it is not true. In this context, the meaning of John's sentence is pragmatically enriched: 'A dog is playing in the garden' is applicable if there is a dog playing in the garden, and it is not Max. The fact that John does not know the dog is a pragmatic inference, not part of the semantic denotation of the sentence.

This example illustrates that, to fully understand a system of communication, it is crucial to posit a division of labour between the semantic denotation of a signal and further pragmatic inferences. This difficulty is particularly problematic for researchers in animal linguistics, who have only access to observations and experiments to derive conclusions about the semantics of a signal, and can draw limited inferences about the pragmatic mechanisms at play in other species. To help in this matter, Schlenker *et al.* (2016*b*) postulated three pragmatic principles that can be applied to any system to unveil the distinction between pragmatics and semantics.

The *informativity principle* postulates that when one signal is strictly more informative than another, the most informative one is used whenever possible. This leads to the assumptions that (*i*) a signaller does not emit a signal S in a situation W if a strictly more informative alternative S′ is applicable in W, and (*ii*) a receiver should infer that if S is emitted, every strictly more informative alternative S′ is non-applicable in W. This was argued to be the case in titi monkeys, whose A-calls refer to serious threats, while B-calls refer to all noteworthy events but are never used for serious threats because A-calls are more appropriate (Commier & Berthet, 2019).

The *urgency principle* postulates that urgent information (e.g. nature or location of a threat) should be communicated as soon as possible in a sequence. As a consequence, signals conveying non-urgent information are used later in the sequence. This principle has been used to explain the use of 'hack' calls of putty-nosed monkeys, which are related to aerial predation situations when used alone or at the start of a sequence, but which convey information about non-ground movements when following other calls (Schlenker *et al.*, 2016*b*,*a*).

These principles are completed with assumptions about the subject's world knowledge. World knowledge is crucial for the receivers to extract information from the utterances. In the example above, Mary is aware that John knows Max, which allows her to draw precise inferences about the

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

meaning of the sentence 'A dog is playing in the garden'. In animals, this knowledge can include the ecology of the species, evaluation of the dangerousness of predators, kinship and affiliative ties among other individuals, *etc.* (see Section II.1.*b*). For example, when hearing calls conveying information about the presence of a serious threat (A-calls), titi monkeys look up, probably because they know that serious threats are raptors (Schlenker *et al.*, 2017; Berthet *et al.*, 2019*a*).

While these principles remain to be experimentally confirmed (but see Narbona Sabaté *et al.*, 2022), they have helped shed light on the vocal systems of several species of monkeys (Schlenker *et al.*, 2016*b*; Dezecache & Berthet, 2018).

(6) Collecting animal data: where to start?

Investigating animal linguistics is challenging because it requires a good understanding of basic linguistic concepts, but also involves collecting data on the behaviour of animals. As such, researchers involved in animal linguistics should be familiar with the basic methodology for collecting and processing animal behavioural data, and be aware of its common pitfalls and difficulties.

Designing and conducting a study with animals requires specific training and knowledge, as well as specific considerations. These aspects have already been extensively covered in the literature, so we will not repeat them here, but we provide a table of references (see Table S1) that may be useful to researchers that are new to the field.

IV. CONCLUSIONS

(1) Animal linguistics is a challenging domain that requires a good knowledge of both linguistics and animal cognition. The study of the evolution of language will benefit from genuine inter-disciplinary collaboration.
(2) One threat is the misapplication of linguistic jargon to animal communication systems. Here, we proposed clear definitions of core concepts in animal linguistics ('semantics', 'pragmatics' and 'syntax'). For each concept, we provide criteria that need to be fulfilled to draw reliable comparisons between human and animal communicative systems.
(3) Another difficulty arises with the choice of relevant and efficient tools to detect linguistic capacities in non-human systems. We reviewed several methods that have already been successfully applied to non-human signals. We hope and expect that additional tools will be developed in future collaborations between linguists and biologists.
(4) A final difficulty comes with the collection of behavioural data on non-human animals. We provide a list of useful references for researchers with little practical knowledge.
(5) This primer aims at encouraging interdisciplinary collaboration, promoting mutual respect among fields and stimulating respectful discussion. We hope it will help the nascent field of animal linguistics thrive and contribute to exciting discoveries on the parallels between animal communication systems and reveal the evolutionary history of language and other communicative systems.

V. ACKNOWLEDGEMENTS

We thank Nathan Klinedinst, Philippe Schlenker and Emmanuel Chemla for their help in the design and preparation of this review. We also thank Alexandra Bosshard, Pritty Patel-Grosz, Patrick Grosz and Dan Sperber for insightful discussions. This research received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 788077, Orisem, PI: Schlenker), the French government IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25), and the British Academy for the Humanities and Social Sciences (grant number PF170023). Research was conducted at Département d'Etudes Cognitives, Ecole Normale Supérieure – PSL Research University, supported by the grant FrontCog ANR-17-EURE-0017, and at the University of Exeter, Center for Ecology and Conservation (Penryn campus). Open access funding provided by Universitat Zurich.

VI. REFERENCES

References identified with an asterisk (*) are cited within the supporting information.

*Adams, D. B. & Kitchen, D. M. (2020). Model vs. playback experiments: the impact of sensory mode on predator-specific escape responses in saki monkeys. *Ethology* **126**, 563–575.

Adams, E. S. & Caldwell, R. L. (1990). Deceptive communication in asymmetric fights of the stomatopod crustacean *Gonodactylus bredini*. *Animal Behaviour* **39**, 706–716.

*Adi, K., Johnson, M. T. & Osiejuk, T. S. (2010). Acoustic censusing using automatic vocalization classification and identity recognition. *The Journal of the Acoustical Society of America* **127**, 874–883.

Alger, S. J., Larget, B. R. & Riters, L. V. (2016). A novel statistical method for behaviour sequence analysis and its application to birdsong. *Animal Behaviour* **116**, 181–193.

Allen, J. A., Garland, E. C., Dunlop, R. A. & Noad, M. J. (2019). Network analysis reveals underlying syntactic features in a vocally learnt mammalian display, humpback whale song. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20192014.

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

*Anusuya, M. A. & Katti, S. K. (2009). Speech recognition by machine: a review. *International Journal of Computer Science and Information Security* **6**, 25.

*Araya-Salas, M. & Smith-Vidaurre, G. (2016). warbleR: an R package to streamline analysis of animal acoustic signals. *Methods in Ecology and Evolution* **8**, 184–191.

Arnold, K. & Bar-On, D. (2020). Primate pragmatics, expressive behavior, and the evolution of language. *Animal Behavior and Cognition* **7**, 117–130.

Arnold, K. & Zuberbühler, K. (2012). Call combinations in monkeys: compositional or idiomatic expressions? *Brain and Language* **120**, 303–309.

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.

*Audacity Team (2014). Audacity. Electronic file available at https://www.audacityteam.org/. Accessed 27.09.2022.

*Avisoft Bioacoustics (2022). Avisoft-SASLab. Electronic file available at https://www.avisoft.com/. Accessed 27.09.2022.

Baayen, H. (1992). Quantitative aspects of morphological productivity. In *Yearbook of Morphology 1991* (eds G. Booij and J. van Marle), pp. 109–149. Springer Netherlands, Dordrecht.

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

*Bearder, S. K., Nekaris, K. A. I., Curtis, D. J., Dew, J. L., Lloyd, J. N. & Setchell, J. M. (2011). Tips from the bush: an A–Z of suggestions for successful fieldwork. In *Field and Laboratory Methods in Primatology: A Practical Guide* (eds J. M. Setchell and D. J. Curtis), pp. 309–323. Cambridge University Press, Cambridge.

Bergman, T. J. & Sheehan, M. J. (2013). Social knowledge and signals in primates. *American Journal of Primatology* **75**, 683–694.

Berthet, M., Benjumea, J., Millet, J., Cäsar, C., Zuberbühler, K. & Dunbar, E. (2019*a*). *Animal Linguistics and the Puzzle of Titi Monkeys Alarm Sequences*. Institute for Logic, Language and Computation, University of Amsterdam, Amsterdam.

Berthet, M., Mesbahi, G., Pajot, A., Cäsar, C., Neumann, C. & Zuberbühler, K. (2019*b*). Titi monkeys combine alarm calls to create probabilistic meaning. *Science Advances* **5**, eaav3991.

Berthet, M., Neumann, C., Mesbahi, G., Cäsar, C. & Zuberbühler, K. (2018). Contextual encoding in titi monkey alarm call sequences. *Behavioral Ecology and Sociobiology* **72**, 8.

Berwick, R. C., Okanoya, K., Beckers, G. J. L. & Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences* **15**, 113–121.

*Bishop, J., Hosey, G. & Plowman, A. (2013). *Handbook of Zoo Research, Guidelines for Conducting Research in Zoos*. British and Irish Association of Zoos and Aquariums, London.

*Boersma, P. & Weenink, D. (2009). Praat: Doing Phonetics by Computer. Electronic file available at www.praat.org/. Accessed 27.09.2022.

Bolhuis, J. J., Beckers, G. J. L., Huybregts, M. A. C., Berwick, R. C. & Everaert, M. B. H. (2018). Meaningful syntactic structure in songbird vocalizations? *PLoS Biology* **16**, e2005157.

Bolinger, D. (1968). Judgments of grammaticality. *Lingua* **21**, 34–40.

Bowling, D. L., Garcia, M., Dunn, J. C., Ruprecht, R., Stewart, A., Frommolt, K. H. & Fitch, W. T. (2017). Body size and vocalization in primates and carnivores. *Scientific Reports* **7**, 41070.

Bradbury, J. W. & Vehrencamp, S. L. (2011). *Principles of Animal Communication*, Second Edition. Sinauer Associates, Sunderland.

*Brainerd, E. L., Blob, R. W., Hedrick, T. L., Creamer, A. T. & Müller, U. K. (2017). Data management rubric for video data in organismal biology. *Integrative and Comparative Biology* **57**, 33–47.

*Brittain, S., Ibbett, H., Lange, E., Dorward, L., Hoyte, S., Marino, A., Milner-Gulland, E. J., Newth, J., Rakotonarivo, S., Veríssimo, D. & Lewis, J. (2020). Ethical considerations when conservation research involves people. *Conservation Biology* **34**, 925–933.

*Brugman, H. & Russel, A. (2004). Annotating multimedia/multi-modal resources with ELAN. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*. European Language Resources Association (ELRA), Lisbon.

*Calisi, R. M. & Bentley, G. E. (2009). Lab and field experiments: are they the same animal? *Hormones and Behavior* **56**, 1–10.

*Center for Conservation Bioacoustics (2010). *Raven pro: Interactive Sound Analysis Software*. The Cornell Lab of Ornithology, Ithaca.

Cheney, D. L. & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences* **115**, 1974–1979.

Chierchia, G. & McConnell-Ginet, S. (1996). *Meaning and Grammar: An Introduction to Semantics*. MIT Press, Cambridge.

Christiansen, M. H. & Kirby, S. (2003). Language evolution: the hardest problem in science? In *Language Evolution* (eds M. H. Christiansen and S. Kirby), pp. 1–15. Oxford University Press, New York, NY.

*Claris (1992). Filemaker Pro. Electronic file available at https://www.claris.com/filemaker/. Accessed 27.09.2022.

*Clink, D.J. & Klinck, H. (2019). GIBBONR: an R package for the detection and classification of acoustic signals using machine learning. Preprint arXiv, 1906.02572. https://doi.org/10.48550/arXiv.1906.02572.

Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B. & Townsend, S. W. (2014). Language evolution: syntax before phonology? *Proceedings of the Royal Society B: Biological Sciences* **281**, 20140263.

Commier, S. & Berthet, M. (2019). Commentary: Titi semantics: context and meaning in titi monkey call sequences. *Frontiers in Psychology* **10**, 512.

Coye, C., Ouattara, K., Zuberbuhler, K. & Lemasson, A. (2015). Suffixation influences receivers' behaviour in non-human primates. *Proceedings of the Royal Society B: Biological Sciences* **282**, 20150265.

Crockford, C., Wittig, R. M., Mundry, R. & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology* **22**, 142–146.

Crockford, C., Wittig, R. M. & Zuberbühler, K. (2017). Vocalizing in chimpanzees is influenced by social-cognitive processes. *Science Advances* **3**, e1701742.

Davidson, G. L., Butler, S., Fernández-Juricic, E., Thornton, A. & Clayton, N. S. (2014). Gaze sensitivity: function and mechanisms from sensory and cognitive perspectives. *Animal Behaviour* **87**, 3–15.

*DDH Software (1997). HanDBase. Electronic file available at http://www.ddhsoftware.com/. Accessed 27.09.2022.

*Deecke, V. B. & Janik, V. M. (2006). Automated categorization of bioacoustic signals: avoiding perceptual pitfalls. *The Journal of the Acoustical Society of America* **119**, 645–653.

Dezecache, G. & Berthet, M. (2018). Working hypotheses on the meaning of general alarm calls. *Animal Behaviour* **142**, 113–118.

*Dian Fossey Gorilla Fund International (2012). Animal Observer. Electronic file available at https://fosseyfund.github.io/AOToolBox/. Accessed 27.09.2022.

Engesser, S., Crane, J. M. S., Savage, J. L., Russell, A. F. & Townsend, S. W. (2015). Experimental evidence for phonemic contrasts in a nonhuman vocal system. *PLoS Biology* **13**, e1002171.

Engesser, S. & Townsend, S. W. (2019). Combinatoriality in the vocal systems of nonhuman animals. *Wiley Interdisciplinary Reviews: Cognitive Science* **10**, e1493.

Ey, E., Pfefferle, D. & Fischer, J. (2007). Do age- and sex-related variations reliably reflect body size in non-human primate vocalizations? A review. *Primates* **48**, 253–267.

*Fedigan, L. M. (2010). Ethical issues faced by field primatologists: asking the relevant questions. *American Journal of Primatology* **72**, 754–771.

*Fischer, J., Noser, R. & Hammerschmidt, K. (2013). Bioacoustic field research: a primer to acoustic analyses and playback experiments with primates. *American Journal of Primatology* **75**, 643–663.

Fitch, W. T. (2005). The evolution of language: a comparative review. *Biology and Philosophy* **20**, 193–203.

Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140091.

Fitch, W. T. (2019). Sequence and hierarchy in vocal rhythms and phonology. *Annals of the New York Academy of Sciences* **1453**, 29–46.

Flower, T. P., Gribble, M. & Ridley, A. R. (2014). Deception by flexible alarm mimicry in an African bird. *Science* **344**, 513–516.

Frege, G. (1892). On concept and object. *Vierteljahrsschrift für wissenschaftliche Philosophie* **16**, 192–205.

Frege, G. (1952). *Translations from the Philosophical Writings of Gottlob Frege*. Blackwell, Oxford.

*Friard, O. & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* **7**, 1325–1330.

*Fukuzawa, Y., Webb, W. H., Pawley, M. D. M., Roper, M. M., Marsland, S., Brunton, D. H. & Gilman, A. (2020). Koe: web-based software to classify acoustic units and analyse sequence structure in animal vocalizations. *Methods in Ecology and Evolution* **11**, 431–441.

*Geissmann, T. & Parson, S. (2011). Recording primate vocalizations. In *Field and Laboratory Methods in Primatology: A Practical Guide*, Second Edition (eds J. M. Setchell and D. J. Curtis), pp. 287–303. Cambridge University Press, Cambridge.

Giannakidou, A. & Zeijlstra, H. (2017). The landscape of negative dependencies: negative concord and N-words. In *The Blackwell Companion to Syntax*, Second Edition (eds M. Everaert and H. van Riemsdijk), pp. 327–391. John Wiley & Sons, Inc., Malden.

Girard-Buttoz, C., Surbeck, M., Samuni, L., Tkaczynski, P., Boesch, C., Fruth, B., Wittig, R. M., Hohmann, G. & Crockford, C. (2020). Information transfer efficiency differs in wild chimpanzees and bonobos, but not social cognition. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20200523.

Girard-Buttoz, C., Zaccarella, E., Bortolato, T., Friederici, A. D., Wittig, R. M. & Crockford, C. (2022). Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology* **5**, 410.

*Goldsmith, M. L. (2005). Habituating primates for field study: ethical considerations for African great apes. In *Biological Anthropology and Ethics: From Repatriation to Genetic Identity* (ed. T. Turner), pp. 49–57. State University of New York Press, New York.

Graham, K. E., Hobaiter, C., Ounsley, J., Furuichi, T. & Byrne, R. W. (2018). Bonobo and chimpanzee gestures overlap extensively in meaning. *PLoS Biology* **16**, e2004825.

*Green, V. M. & Gabriel, K. I. (2020). Researchers' ethical concerns regarding habituating wild-nonhuman primates and perceived ethical duties to their subjects: results of an online survey. *American Journal of Primatology* **82**, e23178.

Grice, P. (1957). Meaning. *The Philosophical Review* **66**, 377–388.

*Hansen, J. H. L. & Hasan, T. (2015). Speaker recognition by machines and humans: a tutorial review. *IEEE Signal Processing Magazine* **32**, 74–99.

*Hanson, K. T. & Riley, E. P. (2018). Beyond neutrality: the human–primate interface during the habituation process. *International Journal of Primatology* **39**, 852–877.

*Hartung, C., Lerer, A., Anokwa, Y., Tseng, C., Brunette, W. & Borriello, G. (2010). Open data kit: tools to build information services for developing regions. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development-ICTD'10*, pp. 1–12. ACM Press, London.

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Hauser, M. D. (1996). Psychological design and communication. In *The Evolution of Communication*, pp. 471–608. The MIT Press, Cambrige; London.

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

Heesen, R., Hobaiter, C., Ferrer-I-Cancho, R. & Semple, S. (2019). Linguistic laws in chimpanzee gestural communication. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20182900.

Hobaiter, C. & Byrne, R. W. (2014). The meanings of chimpanzee gestures. *Current Biology* **24**, 1596–1600.

Hobaiter, C. & Byrne, R. W. (2017). What is a gesture? A meaning-based approach to defining gestural repertoires. *Neuroscience & Biobehavioral Reviews* **82**, 3–12.

*Holle, H. & Rein, R. (2015). EasyDIAg: a tool for easy determination of interrater agreement. *Behavior Research Methods* **47**, 837–847.

*International Primate Society & American Society of Primatologists (2014). Code of best practices for field primatology. Electronic file available at https://internationalprimatologicalsociety.org/wp-content/uploads/2021/10/Code-of_Best_Practices-Oct-2014.pdf. Accessed 27.09.2022.

Janik, V. M. (2000). Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science* **289**, 1355–1357.

Janik, V. M., Sayigh, L. S. & Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences* **103**, 8293–8297.

*Jayne, K. & See, A. (2019). Behavioral research on captive animals: scientific and ethical concerns. In *Animal Experimentation: Working Towards a Paradigm Change* (eds K. Herrmann and K. Jayne), pp. 517–547. BRILL, Leiden, The Netherlands.

Jin, D. Z. & Kozhevnikov, A. A. (2011). A compact statistical model of the song syntax in Bengalese finch. *PLoS Computational Biology* **7**, e1001108.

Kano, F., Krupenye, C., Hirata, S., Tomonaga, M. & Call, J. (2019). Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proceedings of the National Academy of Sciences* **116**, 20904–20909.

Katz, J. J. & Fodor, J. A. (1963). The structure of a semantic theory. *Language* **39**, 170–210.

*Katz, J., Hafner, S. D. & Donovan, T. (2016). Tools for automated acoustic monitoring within the R package monitoR. *Bioacoustics* **25**, 197–210.

*Kaufman, A. B. & Rosenthal, R. (2009). Can you believe my eyes? The importance of interobserver reliability statistics in observations of animal behaviour. *Animal Behaviour* **78**, 1487–1491.

Kean, D., Tiddi, B., Fahy, M., Heistermann, M., Schino, G. & Wheeler, B. C. (2017). Feeling anxious? The mechanisms of vocal deception in tufted capuchin monkeys. *Animal Behaviour* **130**, 37–46.

Kello, C. T., Bella, S. D., Médé, B. & Balasubramaniam, R. (2017). Hierarchical temporal structure in music, speech and animal vocalizations: jazz is like a conversation, humpbacks sing like hermit thrushes. *Journal of the Royal Society Interface* **14**, 20170231.

Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., Bohn, K., Cao, Y., Carter, G., Cäsar, C., Coen, M., DeRuiter, S. L., Doyle, L., Edelman, S., Ferrer-i-Cancho, R., *et al.* (2014*a*). Acoustic sequences in non-human animals: a tutorial review and prospectus. *Biological Reviews* **91**, 13–52.

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

Kershenbaum, A. & Garland, E. C. (2015). Quantifying similarity in animal vocal sequences: which metric performs best? *Methods in Ecology and Evolution* **6**, 1452–1461.

Kershenbaum, A., Ilany, A., Blaustein, L. & Geffen, E. (2012). Syntactic structure and geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society B: Biological Sciences* **279**, 2974–2981.

Kersken, V., Gómez, J.-C., Liszkowski, U., Soldati, A. & Hobaiter, C. (2018). A gestural repertoire of 1- to 2-year-old human children: in search of the ape gestures. *Animal Cognition* **22**, 577–595.

Krupenye, C., Kano, F., Hirata, S., Call, J. & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science* **354**, 110–114.

Kuhn, J., Keenan, S., Arnold, K. & Lemasson, A. (2018). On the *-oo* suffix of Campbell's monkeys. *Linguistic Inquiry* **49**, 169–181.

*Lehner, P. N. (1998). *Handbook of Ethological Methods*, Second Edition. Cambridge University Press, Cambridge, UK.

Leroux, M., Bosshard, A. B., Chandia, B., Manser, A., Zuberbühler, K. & Townsend, S. W. (2021). Chimpanzees combine pant hoots with food calls into larger structures. *Animal Behaviour* **179**, 41–50.

Liao, D. A., Zhang, Y. S., Cai, L. X. & Ghazanfar, A. A. (2018). Internal states and extrinsic factors both determine monkey vocal production. *Proceedings of the National Academy of Sciences* **115**, 3978–3983.

*Liebenberg, L., Steventon, L., Benadie, K. & Minye, J. (1999). Rhino Tracking with the CyberTracker Field Computer. *Pachyderm* **27**, 59–61.

Maille, A., Engelhart, L., Bourjade, M. & Blois-Heulin, C. (2012). To beg, or not to beg? That is the question: Mangabeys modify their production of requesting gestures in response to human's attentional states. *PLoS One* **7**, e41197.

Marantz, A. (2005). Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review* **22**, 429–445.

Markowitz, J. E., Ivie, E., Kligler, L. & Gardner, T. J. (2013). Long-range order in canary song. *PLoS Computational Biology* **9**, e1003052.

Marler, P. (1977). The structure of animal communication sounds. In *Recognition of Complex Acoustic Signals* (eds T. Bullock and E. Evans), pp. 17–35, Dahlem Konferenzen, Berlin.

*Martin, P. R. & Bateson, P. (2007). *Measuring Behaviour: An Introductory Guide*, Third Edition. Cambridge University Press, Cambridge.

*Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W. & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience* **21**, 1281–1289.

*McDonald, M. & Johnson, S. (2014). 'There's an app for that': a new program for the collection of behavioural field data. *Animal Behaviour* **95**, 81–87.

Mehon, F. G. & Stephan, C. (2021). Female putty-nosed monkeys (*Cercopithecus nictitans*) vocally recruit males for predator defence. *Royal Society Open Science* **8**, rsos.202135.

*Melfi, V., Dorey, N. R. & Ward, S. J. (eds) (2020). *Zoo Animal Learning and Training*, First Edition. Wiley-Blackwell, Hoboken.

Mocha, Y. B. & Burkart, J. M. (2021). Intentional communication: solving methodological issues to assigning first-order intentional signalling. *Biological Reviews* **96**, 903–921.

Mumm, C. A. S., Urrutia, M. C. & Knörnschild, M. (2014). Vocal individuality in cohesion calls of giant otters, *Pteronura brasiliensis*. *Animal Behaviour* **88**, 243–252.

*Mundry, R. & Sommer, C. (2007). Discriminant function analysis with nonindependent data: consequences and an alternative. *Animal Behaviour* **74**, 965–976.

Murray, T. G., Zeil, J. & Magrath, R. D. (2017). Sounds of modified flight feathers reliably signal danger in a pigeon. *Current Biology* **27**, 3520–3525.e4.

Narbona Sabaté, L., Mesbahi, G., Dezecache, G., Cäsar, C., Zuberbühler, K. & Berthet, M. (2022). Animal linguistics in the making: the urgency principle and titi monkeys' alarm system. *Ethology Ecology & Evolution* **34**, 378–394.

*Nash, L. T. (2005). Studies of primates in the field and in captivity: similarities and differences in ethical concerns. In *Biological Anthropology and Ethics: From Repatriation to Genetic Identity* (ed. T. Turner), pp. 27–48. State University of New York Press, Albany.

*Newton-Fisher, N.E. (2014). Animal Behaviour Pro. Electronic file available at https://research.kent.ac.uk/lprg/software/. Accessed 27.09.2022.

Olmstead, M. C. & Kuhlmeier, V. A. (2015). *Comparative Cognition*. Cambridge University Press, New York.

*Oswald, J. N., Rankin, S., Barlow, J. & Lammers, M. O. (2007). A tool for real-time acoustic species identification of delphinid whistles. *Journal of the Acoustical Society of America* **122**, 587–595.

Ouattara, K., Lemasson, A. & Zuberbühler, K. (2009*a*). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences* **106**, 22026–22031.

Ouattara, K., Lemasson, A. & Zuberbühler, K. (2009*b*). Campbell's monkeys use affixation to alter call meaning. *PLoS One* **4**, e7808.

Pierce, C. S. (1931). *The Collected Papers of Charles Sanders Pierce*. Harvard University Press, Cambridge.

Prat, Y. (2019). Animals have no language, and humans are animals too. *Perspectives on Psychological Science* **14**, 885–893.

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* **1**, 515–526.

Price, T., Wadewitz, P., Cheney, D., Seyfarth, R., Hammerschmidt, K. & Fischer, J. (2015). Vervets revisited: a quantitative analysis of alarm call structure and context specificity. *Scientific Reports* **5**, 13220.

Pullum, G. K. & Zwicky, A. M. (1988). The syntax-phonology interface. In *Linguistics: The Cambridge Survey* (ed. F. J. Newmeyer), pp. 255–280. Cambridge University Press, Cambridge.

*Riley, E. P. & Bezanson, M. (2018). Ethics of primate fieldwork: toward an ethically engaged primatology. *Annual Review of Anthropology* **47**, 493–512.

*Ross, M.R., Niemann, T., Wark, J.D., Heintz, M.R., Horrigan, A., Cronin, K. A., Shender, M.A. & Gillespie, K. (2016). ZooMonitor. Electronic file available at https://zoomonitor.org/. Accessed 27.09.2022.

*Rowe, N. & Myers, M. (2011). Photography and video for field researchers. In *Field and Laboratory Methods in Primatology: A Practical Guide*, Second Edition (eds J. M. Setchell and D. J. Curtis), pp. 305–318. Cambridge University Press, Cambridge.

*Sainburg, T., Thielk, M. & Gentner, T. Q. (2020). Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. *PLoS Computational Biology* **16**, e1008228.

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Sainburg, T., Theilman, B., Thielk, M. & Gentner, T. Q. (2019). Parallels in the sequential organization of birdsong and human speech. *Nature Communications* **10**, 3636.

Schamberg, I., Wittig, R. M. & Crockford, C. (2018). Call type signals caller goal: a new take on ultimate and proximate influences in vocal production. *Biological Reviews* **93**, 2071–2082.

Schehka, S. & Zimmermann, E. (2009). Acoustic features to arousal and identity in disturbance calls of tree shrews (*Tupaia belangeri*). *Behavioural Brain Research* **203**, 223–231.

Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., Stephan, C., Ryder, R. & Zuberbühler, K. (2014). Monkey semantics: two 'dialects' of Campbell's monkey alarm calls. *Linguistics and Philosophy* **37**, 439–501.

Schlenker, P., Chemla, E., Arnold, K. & Zuberbühler, K. (2016*a*). Pyow-hack revisited: two analyses of putty-nosed monkey alarm calls. *Lingua* **171**, 1–23.

Schlenker, P., Chemla, E., Cäsar, C., Ryder, R. & Zuberbühler, K. (2017). Titi semantics: context and meaning in Titi monkey call sequences. *Natural Language and Linguistic Theory* **35**, 271–298.

Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinović, D., Arnold, K., Cäsar, C., Keenan, S., Lemasson, A., Ouattara, K., Ryder, R. & Zuberbühler, K. (2016*b*). Formal monkey linguistics. *Theoretical Linguistics* **42**, 1–90.

Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinovic, D., Arnold, K., Cäsar, C., Keenan, S., Lemasson, A., Ouattara, K., Ryder, R. & Zuberbühler, K. (2016*c*). Formal monkey linguistics: the debate. *Theoretical Linguistics* **42**, 173–201.

*Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D. & Carvalho, S. (2019). Chimpanzee face recognition from videos in the wild using deep learning. *Science Advances* **5**, eaaw0736.

Scott-Phillips, T. C. (2015*a*). Meaning in animal and human communication. *Animal Cognition* **18**, 801–805.

Scott-Phillips, T. C. (2015*b*). Nonhuman primate communication, pragmatics, and the origins of language. *Current Anthropology* **56**, 56–80.

Scott-Phillips, T. C. (2017). Pragmatics and the aims of language evolution. *Psychonomic Bulletin & Review* **24**, 186–189.

Searcy, W. A. (ed.) (2019). Cognition and language [special issue]. *Animal Behaviour* **151**, 203–274.

*Setchell, J. M. (2019*a*). Conducting fieldwork. In *Studying Primates: How to Design, Conduct and Report Primatological Research*, First Edition, pp. 247–262. Cambridge University Press, Cambridge.

*Setchell, J. M. (2019*b*). *Studying Primates: How to Design, Conduct and Report Primatological Research*. Cambridge University Press, Cambridge.

Shettleworth, S. J. (2010). *Cognition, Evolution, and Behavior*, Second Edition. Oxford University Press, Oxford; New York.

Silk, J. B., Seyfarth, R. M. & Cheney, D. L. (2016). Strategic use of affiliative vocalizations by wild female baboons. *PLoS One* **11**, e0163978.

Slocombe, K. E. & Zuberbuhler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences* **104**, 17228–17233.

*Spillmann, B., van Noordwijk, M. A., Willems, E. P., Mitra Setia, T., Wipfli, U. & van Schaik, C. P. (2015). Validation of an acoustic location system to monitor Bornean orangutan (*Pongo pygmaeus wurmbii*) long calls: triangulation of an acoustic source. *American Journal of Primatology* **77**, 767–776.

*Spillmann, B., van Schaik, C. P., Setia, T. M. & Sadjadi, S. O. (2017). Who shall I say is calling? Validation of a caller recognition procedure in Bornean flanged male orangutan (*Pongo pygmaeus wurmbii*) long calls. *Bioacoustics* **26**, 109–120.

Sprouse, J. (2013). A comparison of informal and formal acceptability judgments using a random sample from *Linguistic Inquiry* 2001–2010. *Lingua* **134**, 219–248.

*Stevenson, J., Liebenberg, L., Derbecker, M., Bapat, V. & Miles, D.G. (1997). CyberTracker. Cape Town.

*Sueur, J. (2018). *Sound Analysis and Synthesis with R*. Springer International Publishing, Cham.

Suzuki, R., Buck, J. R. & Tyack, P. L. (2006). Information entropy of humpback whale songs. *Journal of the Acoustical Society of America* **119**, 1849–1866.

Suzuki, T. N. (2018). Alarm calls evoke a visual search image of a predator in birds. *Proceedings of the National Academy of Sciences* **115**, 1541–1545.

Suzuki, T. N., Wheatcroft, D. & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications* **7**, 10986.

Suzuki, T. N., Wheatcroft, D. & Griesser, M. (2017). Wild birds use an ordering rule to decode novel call sequences. *Current Biology* **27**, 2331–2336.e3.

Suzuki, T. N., Wheatcroft, D. & Griesser, M. (2018). Call combinations in birds and the evolution of compositional syntax. *PLoS Biology* **16**, e2006532.

Suzuki, T. N., Wheatcroft, D. & Griesser, M. (2020). The syntax–semantics interface in animal vocal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**, 20180405.

Szabó, Z. G. (2020). Compositionality. In *The Stanford Encyclopedia of Philosophy* (eds E. N. Zalta and U. Nodelman). Metaphysics Research Lab, Stanford University, Stanfort.

Templeton, C. N., Greene, E. & Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science* **308**, 1934–1937.

*Tuyttens, F. A. M., de Graaf, S., Heerkens, J. L. T., Jacobs, L., Nalon, E., Ott, S., Stadig, L., van Laer, E. & Ampe, B. (2014). Observer bias in animal behaviour research: can we believe what we score, if we score what we believe? *Animal Behaviour* **90**, 273–280.

*Versteegh, M., Kuhn, J., Synnaeve, G., Ravaux, L., Chemla, E., Cäsar, C., Fuller, J., Murphy, D., Schel, A. M. & Dunbar, E. (2016). Classification and automatic transcription of primate calls. *Journal of Acoustical Society of America* **140**, EL26.

*Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F. & Fischer, J. (2015). Characterizing vocal repertoires—hard vs. soft classification approaches. *PLoS One* **10**, e0125785.

*Webster, M. M. & Rutz, C. (2020). How STRANGE are your study animals? *Nature* **582**, 337–340.

Wheeler, B. C. & 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.

*Wildlife Acoustics (2022). Kaleidoscope Analysis Software. Electronic file available at https://www.wildlifeacoustics.com/products/kaleidoscope. Accessed 27.09.2022.

*Williamson, E. A. & Feistner, A. T. (2011). Habituating primates: processes, techniques, variables and ethics. In *Field and Laboratory Methods in Primatology: A Practical Guide* (eds J. M. Setchell and D. J. Curtis), pp. 33–49. Cambridge University Press, Cambridge.

Wittig, R. M., Crockford, C., Langergraber, K. E. & Zuberbühler, K. (2014). Triadic social interactions operate across time: a field experiment with wild chimpanzees. *Proceedings of the Royal Society B: Biological Sciences* **281**, 20133155.

*Wolfe, L. D. (2005). Field primatologists: duties, rights and obligations. In *Biological Anthropology and Ethics: From Repatriation to Genetic Identity* (ed. T. Turner), pp. 15–27. State University of New York Press, Albany.

*Zimmerman, P. H., Bolhuis, J. E., Willemsen, A., Meyer, E. S. & Noldus, L. P. J. J. (2009). The Observer XT: a tool for the integration and synchronization of multimodal signals. *Behavior Research Methods* **41**, 731–735.

*Zollinger, S. A., Podos, J., Nemeth, E., Goller, F. & Brumm, H. (2012). On the relationship between, and measurement of, amplitude and frequency in birdsong. *Animal Behaviour* **84**, e1–e9.

Zuberbühler, K., Cheney, D. & Seyfarth, R. (1999). Conceptual semantic in a nonhuman primate. *Journal of Comparative Psychology* **113**, 33–42.

*Zuberbühler, K. & Wittig, R. M. (2011). Field experiments with non-human primates: a tutorial. In *Field and Laboratory Methods in Primatology: A Practical Guide* (eds J. M. Setchell and D. J. Curtis), pp. 207–224. Cambridge University Press, Cambridge.

VII. SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Useful references to design animal communication studies and collect, process, store and analyse data.

(*Received 5 May 2022; revised 10 August 2022; accepted 12 August 2022*)

1469185x, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/brv.12897 by Ecole Normale Supérieure de Paris, Wiley Online Library on [11/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License