PROCEEDINGS B

Review

Check for updates

Taking turns: bridging the gap between human and animal communication

Simone Pika[1,2], Ray Wilkinson[3], Kobin H. Kendrick[4] and Sonja C. Vernes[5,6]

[1]Department of Primatology, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany
[2]Department of Comparative Biocognition, Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany
[3]Department of Human Communication Sciences, University of Sheffield, Sheffield, UK
[4]Department of Language and Linguistic Science, University of York, York, UK
[5]Neurogenetics of Vocal Communication, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[6]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

SP, 0000-0002-4398-2337; RW, 0000-0002-8036-4415; KHK, 0000-0002-6656-1439; SCV, 0000-0003-0305-4584

Language, humans' most distinctive trait, still remains a 'mystery' for evolutionary theory. It is underpinned by a universal infrastructure—cooperative turn-taking—which has been suggested as an ancient mechanism bridging the existing gap between the articulate human species and their inarticulate primate cousins. However, we know remarkably little about turn-taking systems of non-human animals, and methodological confounds have often prevented meaningful cross-species comparisons. Thus, the extent to which cooperative turn-taking is uniquely human or represents a homologous and/or analogous trait is currently unknown. The present paper draws attention to this promising research avenue by providing an overview of the state of the art of turn-taking in four animal taxa—birds, mammals, insects and anurans. It concludes with a new comparative framework to spur more research into this research domain and to test which elements of the human turn-taking system are shared across species and taxa.

1. Introduction

Language—the most distinctive human trait—remains a 'mystery' [1] or even a 'problem' for evolutionary theory [2,3]. Spoken languages can be characterized by two unique characteristics—a rich learned acoustic portfolio, and the predisposition to combine basic linguistic units into complex acoustic structures [4]. Languages differ at every level of construction, from the sounds, to syntax, to meaning embodying an unrivalled complexity, flexibility and expressivity combined with an unparalleled inter-group variation [5]. Traditionally, comparative studies aiming to unravel the evolutionary trajectory of language have tried to pinpoint the key modalities involved (gestures, vocalizations, combinations of gestures and vocalizations; [6,7]), and/or the underlying complexity in relation to production, usage and comprehension [8]. Recent advances in the fields of Cognitive Sciences, Genetics, Linguistics and Neurosciences, however, suggest that language is a relatively new invention composed of layers of abilities of different types and different antiquity [5,7]. Unpeeling these layers should enable us to understand which distinctive mechanisms were already in place when language first evolved from the communication systems of non-human primates.

In light of this view, an increasing amount of research attention has lately been devoted to the turn-taking system for conversation [9]. It is characterized by a reciprocal exchange of alternating, short and flexible turns between two or more interactants, is used universally across languages and cultures [10], and is based on specific properties [9]. Turn-taking skills develop earlier in ontogeny than gestural and linguistic competence [11], and show some signs of phylogenetic parallels in all clades of the primate lineage [5]. Levinson & Holler ([5], see also [12]) thus proposed that turn-taking

Box 1. Historical definitions of coordinated communicative exchanges

Duetting. This term—sometimes also referred to as dialogue [21,22] or calling songs in insects [23]—traditionally concerns acoustic interactions between two partners of the opposite sex [24]. While some scholars use the term to denote only loosely coordinated behaviours [25], others emphasize the predictable and stereotyped temporal association between an initiating call and its reply [24,26]. Some authors restrict the use of the term to the exact synchronization of identical notes [26], or overlapping bouts of vocalizations [27–29], while others also embrace both airborne and substrate acoustic signals (e.g. antiphonal tapping in woodpeckers), non-acoustic movements (*pas de deux*; [30]) and bioluminescence [24].

Chorusing. This term refers either to a cacophony of sounds or to the synchronous production of the same call type by more than two individuals [31].

Antiphonal singing/song. This term denotes a specialized form of duetting, in which one member of the pair starts a song that is then continued by the other member. The second member may complete the song or the members of the pair may take turns until completion [32].

Antiphonal calling. This term—sometimes also referred to as call-and-response—is defined as a minimum number of two individuals of any sex and/or age combination producing a vocalization in response to a preceding call [31]. Some authors use the term exclusively to denote vocalization exchanges involving the same call type only [33].

Turn-taking. This term was traditionally restricted to human spoken conversation but has recently been extended to other modalities and species. It denotes the orderly exchange of purely communicative signals or behaviours (e.g. peek-a-boo games in humans) between individuals characterized by principles for the coordination of turn transfer, which result in observable temporal regularities. The communicative signals delivered by turns can vary, as can the size and the order of turns, and techniques used to allocate turns to specific individuals [34]. Some scholars see turn-taking as an extension of 'duetting-like' vocal coordination to any conspecific [35].

---

may bridge the apparent gulf between the articulate human species and our non-articulate primate cousins. This hypothesis challenges the predominant view in the field, seeing language as part of a larger uniquely human adaptation for cooperation and cultural life in general [13]. The empirical pillars on which both hypotheses rest are, however, surprisingly weak: we know remarkably little about turn-taking systems of non-human primates and other animals, and methodological confounds have often prevented meaningful cross-species comparisons. Thus, the extent to which cooperative turn-taking is uniquely human or represents a homologous (by shared inheritance) and/or analogous (by parallel evolution) trait is currently unknown. A bias towards purely experimental set-ups and specific communicative modalities, such as the vocal modality, also hampers our understanding of turn-taking systems across different animal species [14,15]. These issues severely impair our understanding of this critical phenomenon, its phylogenetic history and the cognitive underpinnings enabling language to proliferate.

The present paper will draw attention to this promising research avenue by providing an overview of the state of the art of turn-taking in non-human animal communication. As there has been a tremendous number of publications on temporally coordinated signalling—and we can only cite some of them—we have restricted our selection of citations to publications that either cite secondary literature or are exemplary for general phenomena. We will briefly define the predominant terms used, explain the main functional hypotheses proposed, and then discuss their implications for current findings in four animal taxa—birds, mammals, insects and anurans. We conclude that a systematic quantitative comparison of a representative range of turn-taking skills among a single set of human and non-human animal individuals is needed to test the hypotheses of Levinson & Holler [5,12] and Tomasello [13]. To instigate such a comparison, we present a new framework enabling systematic, quantitative assessments of turn-taking abilities across species and taxa. We hope that this framework will spur more quantitative comparative work (and potentially falsify our claims), and shed light on the

question of whether turn-taking has been the 'small change' that made a big difference in human history.

2. Turn-taking and related phenomena

The turn-taking system for conversation [9] applies equally to dyadic and multi-person interactions and is structured and organized according to set principles: alternating, and often relatively short, turns of varying size and order are exchanged between speakers, with only one party normally talking at a time. Speakers construct their turns out of units whose structure allows the next speaker to anticipate their completion. Turn transfer occurs at such points of completion, and turn-allocation works via specific techniques to minimize the temporal gaps between turns ($\approx$200 ms [10]). This communicative exchange is seen as a fundamentally cooperative enterprise [16], involving elements such as 'who should talk or move or act next and when should they do so' ([17], p. 71). It allows interactants to coordinate turn allocation [17], avoid overlap (e.g. [18,19]) and inform others of things helpfully and/or share gossip freely [13], among a myriad of other social actions [20].

In non-human animal studies, multiple and not always mutually exclusive terms have been used to describe coordinated—and not solely communicative—exchanges between interactants involving alternating turns (box 1). Different terminology has been applied to refer to the same phenomenon (e.g. duetting and antiphonal calling) and the same terminology to depict different phenomena (e.g. duetting), while some terms are not mutually exclusive (e.g. turn-taking versus antiphonal calling or duetting). As this paper is concerned with phenomena most closely related to temporally coordinated turn-taking in human conversation, we do not review findings on chorusing, which is lacking in fixed time latencies [31].

3. The function of turn-taking[1]

Although temporal coordination in animal communication has attracted interest over several decades, no clear picture has yet

emerged as to why individuals exchange signals. The explanations put forward generally have not distinguished between the function of signal exchange as such and the function of the exchange of specific signals (e.g. contact calls), though the motivations for these may differ. The earliest hypotheses on the function of duets[2]—inspired by the extraordinary precision of antiphonal singing in tropical bird species—focused on mutual recognition, maintenance of contact between partners, as well as mutual stimulation, reassurance after disturbance and territory defence [36,37]. In the 1980s, these hypotheses were challenged by Wickler ([25], following Armstrong [38]) based on the observed linkage between duetting and monogamy in birds and primates. Wickler tried to explain why temporally coordinated bird songs should be more effective than a solo song. In his view, bird duets function to strengthen the pair-bond by (i) maintaining contact between partners, (ii) synchronizing reproductive physiology, or (iii) advertising mated status. The resulting 'coyness' hypothesis postulated that pair-specific duets are costly strategies, because a high degree of song coordination between pairs is likely to take time and investment. New partners thus need to invest a lot of time learning to duet with the partner, deterring philanderers and making desertion less common [25,30]. In the 1990s, Levin [39] argued that duetting might be a consequence of conflict between the sexes. Recent reviews [27,40] suggest that duets can be multifunctional, including joint-resource defence, signalling commitment, maintaining contact, ensuring reproductive synchrony and mate-guarding. In addition, functions may differ between the sexes, can involve elements of both cooperation and conflict, and/or serve different functions in different circumstances [40,41].

In stark contrast, relatively little research attention has focused on the function of duets in amphibians and insects, and antiphonal exchanges in monogamous and polygamous living societies. The few existing studies on amphibians (*Anura*) and different orders of insects (*Orthoptera, Plecoptera, Hemiptera, Neuroptera*) suggest that the primary function of duets is to enable copulation by acting as mate-location devices [24,42]. Antiphonal exchanges may have partially similar functions as duetting [30], but may also be used to signal social rank and individuality, for coordination, individual recognition, maintenance of social bonds, social cohesion, social integration, and territory defence [25,29,43–48]. Furthermore, virtually nothing is known about the function of turn-taking in human societies [9,12].

4. Turn-taking in birds

Communicative vocal interactions of birds have been intensively studied for more than 50 years. To date, more than 360 species producing vocal duets have been reported [32,40,49]. In this section, we will focus on key aspects and common themes commonly investigated in communicative exchanges of birds including the type of signal used, the time window/temporal relationships (for definition, see Section 9 element (C); [9]), and the avoidance of overlap.

Overall, birds use a large variety of different signals across species, which range from simple calls (e.g. ka-ka; large-billed crow, *Corvus macrorhynchos*) [50] to extended songs (e.g. lesser skylark, *Alauda gulgula*) [51]. Interactions may involve each bird producing the same or a different call/song in response to the initiating vocalization. These interactions take place in different ways, with some species singing the same song in unison [30], while others coordinate their vocal output to produce different components of the same song [52], or engage in countersinging (where a second bird sings a coordinated but overlapping song) [18]. Some bird species, such as nightingales (*Luscinia megarhynchos*), perform 'song matching' (where the bird responds with the same song) or 'vocal supplementing' (where the bird responds with a different, but an appropriate continuation of the initial song) [18].

Research on the time window of avian vocal interactions has focused predominantly on duetting, while investigations into the time window of call exchanges in non-duetting species are relatively rare [50]. Information on the temporal precision of duets is available for 33 species across five orders (galliformes, gruiformes, psittaciformes, piciformes and passeriformes), with most research attention devoted to the order Passeriformes (which includes oscine passerines (songbirds) [40]). Temporal precision in most of these species is relatively high, with latencies between notes ranging from less than 50 ms [53] to 200 ms [40].

Analyses of time-specific relationships within vocal exchanges provided evidence that birds listen and respond to each other and show substantial flexibility in their temporal adjustments [18]. For instance, territorial common nightingales are able to precisely tune their song onset latencies with a peak of approximately 1 s after a neighbour has terminated his song [54]. Results from play-back experiments show that individuals are able to flexibly adjust and shift their latency peaks to account for changes in song duration of stimulus songs and to avoid overlap [54]. The phenomenon of overlap avoidance has been widely documented in several bird species, with the most detailed investigations focusing on nightingales, lesser skylarks [55] and large-billed crows [50]. Avoidance of overlap with regards to development has been studied in barn owls (*Tyto alba*) [56] and European starlings (*Sturnus vulgaris*) [57]. Nestmates of European starlings, for example, exchange calls already very early in ontogeny in the absence of their parents with simultaneous calls occurring below chance level [57]. This finding suggests that distinct time windows may either be learned early, or represent (partially) an innate mechanism (see also the section on non-human primates below).

In addition to avoidance of song overlap, birds have also been observed to adopt two additional roles: they either follow their temporal self-program—called 'autonomous songsters'—or start their songs sometime before (preferentially 1 s after song onset) a neighbour has finished singing—called 'overlappers' [18,54]. This diversity of behaviour represents flexible interaction strategies in some species and species-specific preferences in others. For instance, nightingales adopt different interaction roles in relation to season and social context [18]. By contrast, black-capped chickadees (*Poecile atricapillus*) favour overlap, and European starlings prefer overlap avoidance [58–60]. If overlap occurs, individuals become silent or fly away [60], suggesting that overlapping may be treated, in this species, as a violation of socially accepted rules of turn-taking [60]. It has also been speculated that the overlap itself carries communicative information such as signalling aggression or displaying dominance status [61] or results in direct fitness benefits [62]. For example, a study on quails (e.g. *Lophortyx californicus*) showed that males masking their females' identity prevented other unmated males from mating [62].

In addition, temporal patterns of vocal interactions seem to be tightly linked to a species' social structure [60]. For example, a study on closely related species of Sturnids (African pied starling, *Spreo bicolor*; Cape glossy starling, *Lamprotornis nitens*:

3

red-winged starling, *Onychognathus morio*; pale-winged starling, *Onychognathus nabouroup*) showed that the degree of territoriality highly influenced temporal relationships: the more communal the species, the more song overlap and choruses were observed during close-range interactions [60].

In sum, the use of distinct time windows in birds differs between but also within species ranging from overlap avoidance—representing a characteristic element of human turn-taking [9]—to the strategy of overlapping.

5. Turn-taking in mammals

Research into turn-taking propensities of mammals is strongly biased towards non-human primates. Within the order of primates, studies have nearly exclusively been focusing on vocal exchanges of pair-bonded, and/or family living species (e.g. [28,48,63]). Recently, researchers have also started to investigate turn-taking skills in the gestural modality [34,64], and expanded the research angle onto species living in multi-level [65] and fission–fusion societies [34].

(a) Non-human primates

Some signs of turn-taking have been documented in all the major primate branches [12]: prosimians (e.g. *Lepilemur* spp. [66]; *Tarsius spectrum* [67]), New World monkeys (*Callicebus cupreus* [29]; *Callithrix jacchus* [63]; *Cebuella pygmaea* [48]; *Saimiri* spp. [65]), Old World monkeys (*Cercopithecus campbelli* [68]; *Theropithecus gelada* [69]), smaller apes (*Hylobates* spp. [28,70]) and great apes (*Gorilla gorilla*; *Pan paniscus*; *Pan troglodytes*; *Pongo abelii* [34,64,71]). The vast majority of research has focused on the structure and function of duets in monogamous primate species (e.g. indris *Indri indri*; gibbons *Hylobates* spp.; Mentawi langurs *Presbytis potenziani*; Titi monkeys *Callicebus cupreus* [28,29]), and the antiphonal call exchanges of a distinct clade of New World monkeys, the Callitrichids (e.g. [48,63]). These studies provided evidence that duets are initiated by both communication partners, are pure in tone (i.e. all of the sound energy is compressed into a narrow frequency band), and show manifold diversity [28]. In stark contrast, antiphonal call exchanges are relatively short and most often composed of single call types only (e.g. phee-calls of common marmosets [72]). Similar to antiphonal turn-taking in pair-bonded and/ or family living species, turn-taking in polygynous societies seem to occur mainly between affiliated individuals [65]. Detailed studies on call exchanges of members of the Callitrichid family provide evidence for reciprocal coordination of vocal output [63], and sequential 'conversational' structuring [48]. In addition, studies investigating the temporal relationships underlying turn-taking exchanges showed considerable between-species variability ranging from approximately 500 ms in Saimiri monkeys (*Saimiri sciureus*) [73] to 3000–5000 ms in common marmosets (*Callithrix jacchus*) [63]. Although the development of turn-taking in non-human primates is relatively unexplored, studies on common marmosets imply that the use of antiphonal turn-taking is learned during ontogeny and actively guided by parents. For instance, parents responded differently to overlapping calls of their offspring compared to calls that did not overlap ([5] but see [74]). In addition, parents were more likely to interrupt inappropriate call types produced in response to a specific call (e.g. a twitter in response to a phee-call) than appropriate ones (e.g. phee-call in response to a phee-call) ([75] but see [74]).

Furthermore, spontaneous cooperative turn-taking has been observed in communicative gestural interactions of great apes in both captive and natural environments [34,76]. For instance, focusing on a specific sequential environment—joint travel initiations between mother-infant dyads—Fröhlich and colleagues showed that bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*) establish participation frameworks and adjacency pair-like sequences (for definition, see Section 9 element (D); [9]). Gestural responses can match the temporal relationships observed in human speech (bonobos: 200–1400 ms; chimpanzees: 200–1800 ms) but can also be significantly longer.

(b) Non-primate mammals

Outside of the primate order, vocal turn-taking has been studied in four distantly related mammalian groups, cetaceans, bats, elephants and mole rats.

The most research attention has been devoted to cetaceans and provided evidence that vocal exchanges facilitating social interactions occur in a number of species including beluga whales (*Delphinapterus leucas*) [77], bottlenose dolphins (*Tursiops truncatus*) [78], killer wales (*Orcinus orca*) [79], southern right whales (*Eubalaena austrialis*) [80] and sperm whales (*Physeter macrocephalus*) [81]. Bottlenose dolphins represent the best studied cetacean species, partly facilitated by their relatively small size and relative frequency in captivity. Dolphins produce characteristic signature whistles that are used in coordinated vocal interactions and seem to facilitate individual recognition and maintenance of group cohesion. Isolated dolphins in captivity use alternating whistles with minimal overlap when two or more dolphins can interact (physically or only acoustically) [78,82–84] (although in the wild, overlap appears more frequent, see [85]). Time windows of vocal turns between interacting individuals were generally less than 1 s [47,83,84]. Observations on dolphins in their natural environments showed that exchanges of whistles commonly precede an animal joining a group [86]. Similarly, southern right whales exchange a specific call type—the 'up' call—during approach and integration into a group. Pairs of dolphins in captivity have also been observed to partake in duets characterized by closely matched frequencies and timing of whistles. They are also able to swap between alternating (antiphony) and duetting within the same train of vocalizations [78]. Similar to the signature whistles of dolphins, beluga whales use burst pulse sounds. These calls are predominantly produced within a time window of approximately 1 s following a burst pulse sound produced by a conspecific [77]. They thus mirror the timing of signature whistle exchanges in dolphins. By contrast, and possibly due to living in stable rather than fission–fusion societies, beluga whales, killer whales and sperm whales exchange group-specific calls (characteristic for single groups; [79,81,87]). These calls are used in an antiphonal manner and are much more likely to occur within a time window of approximately 5 s [79]. Furthermore, call types are frequently matched (responding to the first call with the same call type) by conspecifics [79], suggesting that the vocal behaviour of group members highly impacts upon timing and call type choice. Similarly, sperm whales exchange sequences of broadband clicks (codas) with either temporal gaps of 2 s or by using overlapping codas, and also show call matching of the original coda type [81].

Much less information is available regarding turn-taking in bats, elephants and mole rats. Bats engage in antiphonal calling between adults (white-winged vampire bats;

Diaemus youngi) [43] or between mothers and their offspring (young pups not capable of flight) in the families of Molossidae, Vespertilionidae, Phyllostomidae and Emballonuridae [44,88–92]. Temporal relationships have so far only been studied in the white-winged vampire bats, ranging from 300–350 ms [43]. Female elephants (*Loxodonta africana*) exchange vocalizations, such as low-frequency rumbles, to respond to calls from other females [93,94]. A response is most likely when the interacting females have strong social relationships [93,94], with call exchanges often resulting in closer proximity between participating animals [46]. Naked mole rats (*Heterocephalus glaber*) are one of the few eusocial mammal species, and use their most common vocalization type—the soft chirp—antiphonally with a latency of approximately 400 ms [45,95].

6. Turn-taking in insects

Research into communicative exchanges of insects has focused on five different orders: the Choleoptera, the Hemiptera, the Neuroptera, the Ortoptera and the Plecoptera [22–24,96]. The signal producing mechanisms are very diverse and range from vibration, percussion, stridulation, over click mechanisms, air expulsion, to bioluminescence [22,97]. The first signal of a given interaction is—in contrast to the more flexible duets of birds and mammals [98]—always initiated by males, with mechanisms often differing consistently across the sexes [24]. For instance, the sounds or substrate vibrations of males of many homopteran cicadellids and cicadas are produced by a tymbal, while the females respond by using vibrations created by movements of the wings [24]. The length of the initiating call is highly variable within-species, while the distinctive temporal pattern and the time window between signal and female reply are often species-specific [99]. Males of a species that initiate duets via long complex calls often insert a trigger pulse at the end of the call [100], which may act as a cue for the female to reply [24]. The variability of signal interaction between the sexes is manifold, ranging from brief exchanges (e.g. stonefly, *Eucoptura xanthenes*) to relatively complex sequences involving females alternating their replies between the pulsed phrases of the male signal (e.g. North American katydid, *Amblycorypha parvipenni*) [24]. Temporal relationships vary from extremely short intervals (e.g. 15 ms, blackwinged saw bush-cricket *Ancistrus nigrovittata*; 20–30 ms, speckled bush-cricket *Leptophyes punctatissima*) to even 850 ms in species relying on bioluminescent systems (e.g. Photinus firefly *Photinus greeni* [96]).

7. Turn-taking in anurans

In anurans, turn-taking mainly takes place in the form of antiphonal advertisement calls by males to attract females and has been observed in most groups of frogs. Males producing vocalizations in close distance to each other (or in response to playbacks) will typically become temporally entrained such that overlap is avoided. Calls thus occur within defined time points after the completion of another call [101–103]. For example, in green frogs (*Rana clamitans*) calls from different males are spaced at intervals of 2–10 s and rarely overlap [104]. Males of the Sri Lankan tree frog (*Philautus leucorhinus*) engage in vocal matching which appears to be based on the nature of the rival's call rather than simply being an example of vocal stereotypy [105]. A small number of species has been

observed where overlap is typical [103]. For example, American toad males (*Bufo americanus*) engage in synchronous, or near-synchronous, overlapping of calls [103,104]. And, while unusual, males and females of some species also engage in duets. In American clawed frogs (*Xenopus laevis*), for example, females may produce a 'rapping' vocalization in response to male calling which is then responded to with a male answer call [106].

8. Conclusion

Overall, direct comparisons of turn-taking skills of non-human animals in relation to language origins (but also social communication and communication in general) are highly constrained by lack of data, the application of different terms, methodological designs (observational versus experimental paradigms) and study environments (captivity versus natural environments). Furthermore, investigations have so far mainly been focusing on single call types (e.g. phee-call in common marmosets) or songs (e.g. great-calls in gibbons) of species, limiting an in-depth understanding of the variability and underlying cognitive flexibility of turn-taking systems found in the animal kingdom.

To date, the parameters tested across different taxa and species to infer the organization of non-human animal turn-taking have mainly concerned a single key element of full-blown human turn-taking—the time window. Hence, the temporal adaptation and alteration of signal production seems to be a basic element of sociality and communication in general, and may have been the first step in the evolution of turn-taking systems. However, it is currently virtually impossible to evaluate whether time windows across species and taxa are indeed similar phenomena or differ because different definitions, methodologies and signal types (differing in form and function) have been investigated.

Most progress concerning an in-depth understanding of the degree of similarity between human and non-human communicative turn-taking systems has been made by studies taking into consideration Sacks *et al.*'s [9] systematics for the organization of human turn-taking. For instance, Rossano [76] and Fröhlich *et al.* [34] investigated gestural interactions of great apes with a special focus on turn-allocation techniques, distinct time windows and adjacency pair-like structures. They showed that bonobos and chimpanzees use gaze and distinct proximity patterns to allocate turns. In addition, both species have species-specific time windows, and are able to form adjacency pair-like structures (e.g. a carrying request resulting in being carried).

Future studies should push this approach even further by testing whether the most crucial hallmarks of human conversational turn-taking can be found in turn-taking systems of other animals: who should communicate, move or act and when should interactants do so [17]. Such an unprecedented rigorous test of turn-taking skills will enable us to gain insight into the layers shared across species and taxa and the cognitive complexity underlying specific elements of the turn-taking system. For instance, rhythmic signalling in many insect and anuran species is controlled by a central nervous system oscillator that may be inhibited, and reset by an acoustic stimulus such as a competitor's call ([103,104,107], but see [108]). There is no response *per se* to preceding signals, while in contrast species-specific time windows in monkey species are learnt [68], with individuals taking into consideration sex, age and rank of recipients [75,109–111].

5

6

rspb.royalsocietypublishing.org   Proc. R. Soc. B 285: 20180598

9. The comparative turn-taking framework

The new framework enabling comparative, systematic, quantitative assessments of turn-taking abilities centres on four key elements characterizing human social action during conversation:

(A) Flexibility of turn-taking organization
(B) Who is taking the next turn?
(C) When do response turns occur?
(D) What should the next turn do?

The first element—flexibility of turn-taking organization (A)—refers to the phenomena of varying size and ordering of turns and intentionality involved in human turn-taking sequences [9]. The element mirrors the ability to voluntarily change and adjust signals/actions and thus the degree of underlying cognitive flexibility. It can be operationalized by quantifying the number, frequency and degree of repetition of signals and actions produced in turn-taking events, their combination (e.g. A-B-A; A-B-C), distribution of roles between participants (e.g. role reversal), and intentionality involved (e.g. goal persistence, sensitivity to the social context) [34,112,113].

The second element—who is taking the next turn (B)—concerns who can or should produce the next signal and includes techniques for allocating turns to individuals or parties [9]. Parameters should involve (i) body orientation towards recipient(s), (ii) gaze direction of signaller, (iii) response waiting, and (iv) whether recipient(s) can perceive the signal (e.g. being in the visual or auditory field).

The third element—when do response turns occur (C)—addresses the *time window* or *temporal relationship* between an initiating turn and the response turn [10,24]. Since the normative timing of signal exchanges may differ across species, modalities, and transmission medium, a first mandatory step should be to establish typical time windows for a given species (see [34] for ideas to operationlize this element).

The fourth element—what should the next turn do? (D)—concerns one of the most fundamental structures in the organization of human conversation: *adjacency pairs* [114]. An adjacency pair can be recursively reproduced [115] and expanded in conversation and—in its minimal, unexpanded form—is composed of two turns, by different participants, that are adjacently placed, and are relatively ordered into *first pair parts* (actions that initiate some exchange, e.g. requests), and *second pair parts* (responsive actions, e.g. grants) [114]. This element can be operationalized by testing whether subsequent turns qualify as adjacency pairs involving predictable signal-response sequences (e.g. a request gesture is typically responded with a granting signal; a call is typically responded with the same call type, e.g. common marmosets) [74,116].

10. Empirical desiderata

A major avenue of future research is to use the comparative turn-taking framework to characterize the turn-taking phenotype of a wide variety of primate species. This could be done through systematic testing of carefully chosen representatives of more than 50 genera of primates, which should then enable us to map out cladistically the evolution of primates' turn-taking skills and systems. Furthermore, recent findings on language competence and cognitive skills of members of the parrot and corvid family [117–119] have put into question the assumed simple inverse correlation between language-readiness and genetic distance from humans. Although avian and primate brains differ significantly in size, structure, and neuron numbers, similar principles of organization are evident, reflecting a case of convergent evolution in relation to mental processes [120,121]. Examples of convergent evolution in distant-related species can, therefore, provide important clues to the types of problems that particular morphological or behavioural mechanisms are 'designed' to solve. Furthermore, in order to claim that particular components of human language are unique to humans, data indicating that no other animal has this particular trait is required.

Such an unprecedented, systematic comparative approach will empower us to test whether cooperative turn-taking represents the most ancient infrastructure of the language system and has been the 'small change' that made a big difference in human history. This new field of comparative turn-taking will thus shed light on one of the 'hardest' problems in science [3] by testing whether turn-taking had profound downstream effects on human culture and cooperation, and laid the foundation for the evolution of language.

Endnotes

[1]From here on, the term 'turn-taking' refers to all temporal coordinated phenomena introduced.
[2]From here on, the terms 'duet' and 'antiphonal' mirror the use of these terms in the reviewed studies, but do not reflect an assessment of the authors.

References

1. Knight C, Studdert-Kennedy M, Hurford JR. 2000 Language: a Darwinian adaptation? In *The evolutionary emergence of language: social function and the origins of linguistic form* (eds C Knight, M Studdert-Kennedy, JR Hurford),

pp. 1–15. Cambridge, UK: Cambridge University Press.
2. Hauser MD, Yang C, Berwick RC, Tattersall I, Ryan MJ, Watumull J, Chomsky N, Lewontin RC. 2014 The mystery of language evolution.

Front. Psychol. **5**, 12. (doi:10.3389/fpsyg.2014.00401)
3. Christiansen MH, Kirby S. 2003 Language evolution: the hardest problem in science? In *Language evolution: the states of the Art* (eds MH Christiansen,

S Kirby), pp. 1–15. Oxford, UK: Oxford University Press.

4. Hauser MD, Chomsky N, Fitch WT. 2002 The faculty of language: what is it, who has it, and how did it evolve? *Science* **298**, 1568–1579. (doi:10.1126/science.298.5598.1569)

5. Levinson SC, Holler J. 2014 The origin of human multi-modal communication. *Phil. Trans. R. Soc. B* **369**, 20130302. (doi:10.1098/rstb.2013.0302)

6. Arbib MA, Liebal K, Pika S. 2008 Primate vocalization, gesture, and the evolution of human language. *Curr. Anthropol.* **49**, 1053–1076. (doi:10.1086/593015)

7. Fitch WT. 2010 *The evolution of language*, p. 624. Cambridge, UK: Cambridge University Press.

8. Seyfarth RM, Cheney DL. 2010 Production, usage, and comprehension in animal vocalizations. *Brain Lang.* **115**, 92–100. (doi:10.1016/j.bandl.2009.10.003)

9. Sacks H, Schegloff EA, Jefferson G. 1974 A simplest systematics for the organization of turn-taking in conversation. *Language* **50**, 696–735. (doi:10.1353/lan.1974.0010)

10. Stivers T et al. 2009 Universals and cultural variation in turn-taking in conversation. *Proc. Natl Acad. Sci. USA* **106**, 10 587–10 592. (doi:10.1073/pnas.0903616106)

11. Bateson MC. 1975 Mother-infant exchanges: the epigenesis of conversational interaction. *Ann. NY Acad. Sci.* **263**, 101–113. (doi:10.1111/j.1749-6632.1975.tb41575.x)

12. Levinson SC. 2016 Turn-taking in human communication—origins and implications for language processing. *Trends Cogn. Sci.* **20**, 6–14. (doi:10.1016/j.tics.2015.10.010)

13. Tomasello M. 2008 *Origins of human communications*. Cambridge, MA: MIT Press.

14. Naguib M, Mennill DJ. 2010 The signal value of birdsong: empirical evidence suggests song overlapping is a signal. *Anim. Behav.* **80**, e11–e15. (doi:10.1016/j.anbehav.2010.06.001)

15. Liebal K, Waller B, Burrows A, Slocombe K. 2013 *Primate communication: a multimodal approach*, p. 306. Cambridge, UK: Cambridge University Press.

16. Grice HP. 1957 Meaning. *Philos. Rev.* **66**, 377–388. (doi:10.2307/2182440)

17. Schegloff EA. 2006 Interaction: the infrastructure for social institutions, the natural ecological niche for language, and the arena in which culture is enacted. In *Roots of human sociality, culture, cognition and interaction* (eds NJ Enfield, SC Levinson), pp. 70–96. Oxford, UK: New York, Berg.

18. Hultsch H, Todt D. 1982 Temporal performance roles during vocal interactions in nightingales (*Luscinia megarhynchos* B.). *Behav. Ecol. Sociobiol.* **11**, 253–260. (doi:10.1007/BF00299302)

19. Naguib M. 1999 Effects of song overlapping and alternating on nocturnally singing nightingales. *Anim. Behav.* **58**, 1061–1067. (doi:10.1006/anbe.1999.1223)

20. Levinson SC. 2013 Action formation and ascription. In *The handbook of conversation analysis* (eds T Stivers, J. Sidnell), pp. 103–130. Malden, MA: Wiley-Blackwell.

21. Buck J, Case J. 2002 Physiological links in firefly flash code evolution. *J. Insect Behav.* **15**, 51–68. (doi:10.1023/A:1014480028731)

22. Greenfield MD. 1994 Cooperation and conflict in the evolution of signal interactions. *Annu. Rev. Ecol. Syst.* **25**, 97–126. (doi:10.1146/annurev.es.25.110194.000525)

23. Drosopoulos S, Claridge MF. 2005 *Insect sounds and communication: physiology, behaviour, ecology, and evolution*, p. 532. London, NY: CRC Press, Taylor & Francis Group.

24. Bailey WJ. 2003 Insect duets: Underlying mechanisms and their evolution. *Physiol. Entomol.* **28**, 157–174. (doi:10.1046/j.1365-3032.2003.00337.x)

25. Wickler W. 1980 Vocal dueting and the pair bond. I. Coyness and partner commitment. A hypothesis. *Zeitschrift für Tierpsychologie* **52**, 201–209. (doi:10.1111/j.1439-0310.1980.tb00711.x)

26. Langmore NE. 2002 Vocal duetting: definitions, discoveries and directions. *Trends Ecol. Evol.* **17**, 451–452. (doi:10.1016/S0169-5347(02)02611-3)

27. Hall ML. 2004 A review of hypotheses for the functions of avian duetting. *Behav. Ecol. Sociobiol.* **55**, 415–430. (doi:10.1007/s00265-003-0741-x)

28. Haimoff EH. 1986 Convergence in the duetting of monogamous old world primates. *J. Hum. Evol.* **15**, 51–59. (doi:10.1016/S0047-2484(86)80065-3)

29. Müller AE, Anzenberger G. 2002 Duetting in the titi monkey *Callicebus cupreus*: Structure, pair specificity and development of duets. *Folia Primatol.* **73**, 104–115. (doi:10.1159/000064788)

30. Wickler W, Seibt U. 1982 Dueting and the pair bond. *Animal Behaviour* **30**, 943–944.

31. Yoshida Y, Okanoya K. 2005 Evolution of turn-taking: a bio-cognitive perspective. *Cogn. Studies* **12**, 153–165. (doi:10.1163/1568539X-00003031)

32. Thorpe WH. 1975 The biological significance of duetting and antiphonal song. *Acta Neurobiol. Exp. (Wars)* **35**, 517–528.

33. Chen HC, Kaplan G, Rogers LJ. 2009 Contact calls of common marmosets (*Callithrix jacchus*): influence of age of caller on antiphonal calling and other vocal responses. *Am. J. Primatol.* **71**, 165–170. (doi:10.1002/ajp.20636)

34. Fröhlich M, Kuchenbuch P, Müller G, Fruth B, Furuichi T, Wittig RM, Pika S. 2016 Unpeeling the layers of language: bonobos and chimpanzees engage in cooperative turn-taking sequences. *Sci. Rep.* **6**, 25887. (doi:10.1038/srep25887)

35. Ghazanfar AA, Takahashi DY. 2014 The evolution of speech: vision, rhythm, cooperation. *Trends Cogn. Sci.* **18**, 543–553. (doi:10.1016/j.tics.2014.06.004)

36. Thorpe WH, Hall-Craggs J, Hooker B, Hooker T, Hutchison R. 1972 *Duetting and antiphonal song in birds: its extent and significance*. Leiden, The Netherlands: E.J. Brill.

37. Snowdon CT. 1989 Vocal communication in New World monkeys. *J. Hum. Evol.* **18**, 611–633. (doi:10.1016/0047-2484(89)90097-3)

38. Armstrong EA. 1963 *A study of bird song*, p. 335. London, Beer, JV.

39. Levin RN. 1996 Song behaviour and reproductive strategies in a duetting wren, *Thryothorus nigricapillus*: II. Playback experiments. *Anim. Behav.* **52**, 1107–1117. (doi:10.1006/anbe.1996.0258)

40. Dahlin CR, Benedict L. 2014 Angry birds need not apply: a perspective on the flexible form and multifunctionality of avian vocal duets. *Ethology* **120**, 1–10. (doi:10.1111/eth.12182)

41. Cowlishaw G. 1992 Song function in gibbons. *Behaviour* **121**, 132–153. (doi:10.1163/156853992X00471)

42. Shen J-X, Feng AS, Xu Z-M, Yu Z-L, Arch VS, Yu X-J, Narins PM. 2008 Ultrasonic frogs show hyperacute phonotaxis to female courtship calls. *Nature* **453**, 914–916. (doi:10.1038/nature06719)

43. Carter GG, Skowronski MD, Faure PA, Fenton B. 2008 Antiphonal calling allows individual discrimination in white-winged vampire bats. *Anim. Behav.* **76**, 1343–1355. (doi:10.1016/j.anbehav.2008.04.023)

44. Knörnschild M, von Helversen O. 2008 Nonmutual vocal mother–pup recognition in the greater sac-winged bat. *Anim. Behav.* **76**, 1001–1009. (doi:10.1016/j.anbehav.2008.05.018)

45. Yosida S, Kobayasi KI, Ikebuchi M, Ozaki R, Okanoya K. 2007 Antiphonal vocalization of a subterranean rodent, the naked mole-rat (*Heterocephalus glaber*). *Ethology* **113**, 703–710. (doi:10.1111/j.1439-0310.2007.01371.x)

46. Leighty KA, Soltis J, Wesolek CM, Savage A. 2008 Rumble vocalizations mediate interpartner distance in African elephants, *Loxodonta africana*. *Anim. Behav.* **76**, 1601–1608. (doi:10.1016/j.anbehav.2008.06.022)

47. Janik VM. 2000 Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science* **289**, 1355–1357. (doi:10.1126/science.289.5483.1355)

48. Snowdon CT, Cleveland J. 1984 'Conversations' among pygmy marmosets. *Am. J. Primatol.* **7**, 15–20. (doi:10.1002/ajp.1350070104)

49. Thorpe WH, Hall-Craggs J, Hooker B, Hooker T, Hutchison R. 1972 Duetting and antiphonal song in birds: its extent and significance. *Behaviour Supplement*, No. 18, 1–197.

50. Kondo N, Watanabe S, Izawa E-I. 2010 A temporal rule in vocal exchange among large-billed crows *Corvus macrorhynchos* in Japan. *Ornithol. Sci.* **9**, 83–91. (doi:10.2326/osj.9.83)

51. Gochfeld M. 1978 Intraspecific social stimulation and temporal displacement of songs of the lesser skylark, *Alauda gulgula*. *Ethology* **48**, 337–344.

52. Catchpole CK, Slater PJB. 1995 *Bird song: biological themes and variation*, p. 256. Cambridge, UK: Cambridge University Press.

53. Mann NI, Dingess KA, Slater PJB. 2006 Antiphonal four-part synchronized chorusing in a Neotropical wren. *Biol. Lett.* **2**, 1–4. (doi:10.1098/rsbl.2005.0373)

54. Todt D, Hultsch H. 1999 How Nightingales develop their vocal competence. In *22. International*

7

rspb.royalsocietypublishing.org *Proc. R. Soc. B* **285**: 20180598

ornithological congress (eds NJ Adams, RH Slotow), pp. 193–215. Durban, South Africa: BirdLife South.

55. Gochfeld M. 2010 Intraspecific social stimulation and temporal displacement of songs of the lesser skylark, *Alauda gulgula*. *Zeitschrift für Tierpsychologie* **48**, 337–344. (doi:10.1111/j.1439-0310.1978.tb00264.x)

56. Dreiss AN, Ruppli CA, Oberli F, Antoniazza S, Henry I, Roulin A. 2013 Barn owls do not interrupt their siblings. *Anim. Behav.* **86**, 119–126. (doi:10.1016/j.anbehav.2013.04.019)

57. Chaiken M. 1990 The ontogeny of antiphonal calling in European starlings. *Dev. Psychobiol.* **23**, 233–246. (doi:10.1002/dev.420230304)

58. Fitzsimmons LP, Foote JR, Ratcliffe LM, Mennill DJ. 2008 Frequency matching, overlapping and movement behaviour in diurnal countersinging interactions of black-capped chickadees. *Anim. Behav.* **75**, 1913–1920. (doi:10.1016/j.anbehav.2007.11.006)

59. Foote JR, Fitzsimmons LP, Mennill DJ, Ratcliffe LM. 2008 Male chickadees match neighbors interactively at dawn: support for the social dynamics hypothesis. *Behav. Ecol.* **19**, 1192–1199. (doi:10.1093/beheco/arn087)

60. Henry L, Craig AJFK, Lemasson A, Hausberger M. 2015 Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Language Sci.* **6**, 1416. (doi:10.3389/fpsyg.2015.01416)

61. Dabelsteen T, McGregor PK, Holland JO, Tobias JA, Pedersen SB. 1997 The signal function of overlapping singing in male robins. *Anim. Behav.* **53**, 249–256. (doi:10.1006/anbe.1996.0369)

62. Stokes AW. 1968 Antiphonal calling in quail. *Auk* **85**, 83–89. (doi:10.2307/4083626)

63. Takahashi DY, Narayanan DZ, Ghazanfar AA. 2013 Coupled oscillator dynamics of vocal turn-taking in monkeys. *Curr. Biol.* **23**, 2162–2168. (doi:10.1016/j.cub.2013.09.005)

64. Rossano F, Liebal K. 2014 'Requests' and 'offers' in orangutans and human infants. In *Requesting in social interaction* (eds P Drew, E Couper-Kuhlen), pp. 335–364. Amsterdam, The Netherlands: John Benjamins Publishing Co.

65. Symmes D, Biben M. 1988 Conversational vocal exchanges in squirrel monkeys. In *Primate vocal communication* (eds PDD Todt, DP Goedeking, DD Symmes), pp. 123–132. Berlin, Germany: Springer.

66. Méndez-Cárdenas MG, Zimmermann E. 2009 Duetting—a mechanism to strengthen pair bonds in a dispersed pair-living primate (*Lepilemur edwardsi*)? *Am. J. Phys. Anthropol.* **139**, 523–532. (doi:10.1002/ajpa.21017)

67. MacKinnon J, MacKinnon K. 1980 The behavior of wild spectral tarsiers. *Int. J. Primatol.* **1**, 361–379. (doi:10.1007/BF02692280)

68. Lemasson A, Glas L, Barbu S, Lacroix A, Guilloux M, Remeuf K, Koda H. 2011 Youngsters do not pay attention to conversational rules: is this so for nonhuman primates? *Sci. Rep.* **1**, 4. (doi:10.1038/srep00022)

69. Richman B. 1976 Some vocal distinctive features used by gelada monkeys. *J. Acoust. Soc. Am.* **60**, 718–724. (doi:10.1121/1.381144)

70. Geissmann T. 1999 Duet songs of the siamang, *Hylobates syndactylus*: II. Testing the pair-bonding hypothesis during a partner exchange. *Behaviour* **136**, 1005–1039. (doi:10.1163/156853999501694)

71. Seyfarth RM, Cheney DL, Harcourt AH, Stewart KJ. 1994 The acoustic features of gorilla double grunts and their relation to behavior. *Am. J. Primatol.* **33**, 31–50. (doi:10.1002/ajp.1350330104)

72. Miller CT, Mandel K, Wang X. 2010 The communicative content of the common marmoset phee call during antiphonal calling. *Am. J. Primatol.* **72**, 974–980. (doi:10.1002/ajp.20854)

73. Masataka N, Biben M. 1987 Temporal rules regulating affiliative vocal exchanges of squirrel monkeys. *Behaviour* **101**, 311–319. (doi:10.1163/156853987X00035)

74. Takahashi DY, Fenley AR, Ghazanfar AA. 2016 Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Phil. Trans. R. Soc. B* **371**, 20150370. (doi:10.1098/rstb.2015.0370)

75. Chow CP, Mitchell JF, Miller CT. 2015 Vocal turn-taking in a non-human primate is learned during ontogeny. *Proc. R. Soc. B* **282**, 20150069. (doi:10.1098/rspb.2015.0069)

76. Rossano F. 2013 Sequence organization and timing of bonobo mother-infant interactions. *Interac. Stud.* **14**, 160–189. (doi:10.1075/is.14.2.02ros)

77. Morisaka T, Yoshida Y, Akune Y, Mishima H, Nishimoto S. 2013 Exchange of 'signature' calls in captive belugas (*Delphinapterus leucas*). *J. Ethol.* **31**, 141–149. (doi:10.1007/s10164-013-0358-0)

78. Lilly JC, Miller AM. 1961 Sounds emitted by the bottlenose dolphin. *Science* **133**, 1689–1693. (doi:10.1126/science.133.3465.1689)

79. Miller PJO, Shapiro AD, Tyack PL, Solow AR. 2004 Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*. *Anim. Behav.* **67**, 1099–1107. (doi:10.1016/j.anbehav.2003.06.017)

80. Clark CW, Clark JM. 1980 Sound playback experiments with southern right whales (*Eubalaena-Australis*). *Science* **207**, 663–665. (doi:10.1126/science.207.4431.663)

81. Schulz TM, Whitehead H, Gero S, Rendell L. 2008 Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Anim. Behav.* **76**, 1977–1988. (doi:10.1016/j.anbehav.2008.07.032)

82. Lang TG, Smith HA. 1965 Communication between dolphins in separate tanks by way of an electronic acoustic link. *Science* **150**, 1839–1844. (doi:10.1126/science.150.3705.1839)

83. Caldwell CA. 1968 Vocalisation of native dolphins in small groups. *Science* **159**(3819), 1121–1123.

84. Nakahara F, Miyazaki N. 2011 Vocal exchanges of signature whistles in bottlenose dolphins (*Tursiops truncatus*). *J. Ethol.* **29**, 309–320. (doi:10.1007/s10164-010-0259-4)

85. Quick NJ, Janik VM. 2012 Bottlenose dolphins exchange signature whistles when meeting at sea. *Proc. R. Soc. B* **279**, 2539–2545. (doi:10.1098/rspb.2011.2537)

86. Janik VM. 2014 Cetacean vocal learning and communication. *Curr. Opin. Neurobiol.* **28**, 60–65. (doi:10.1016/j.conb.2014.06.010)

87. Schulz TM, Whitehead H, Gero S, Rendell L. 2011 Individual vocal production in a sperm whale (*Physeter macrocephalus*) social unit. *Mar. Mamm. Sci.* **27**, 149–166. (doi:10.1111/j.1748-7692.2010.00399.x)

88. Balcombe JP. 1990 Vocal recognition of pups by mother Mexican free-tailed bats, tadarida-brasiliensis-mexicana. *Anim. Behav.* **39**, 960–966. (doi:10.1016/S0003-3472(05)80961-3)

89. Balcombe JP, Mccracken GF. 1992 Vocal recognition in Mexican free-tailed bats—do pups recognize mothers. *Anim. Behav.* **43**, 79–87. (doi:10.1016/S0003-3472(05)80073-9)

90. deFanis E, Jones G. 1996 Allomaternal care and recognition between mothers and young in pipistrelle bats (*Pipistrellus pipistrellus*). *J. Zool.* **240**, 781–787. (doi:10.1111/j.1469-7998.1996.tb05324.x)

91. Esser KH, Schmidt U. 1989 Mother-infant communication in the lesser spear-nosed bat phyllostomus-discolor (Chiroptera, Phyllostomidae)—evidence for acoustic learning. *Ethology* **82**, 156–168. (doi:10.1111/j.1439-0310.1989.tb00496.x)

92. Carter GG, Fenton MB, Faure PA. 2009 White-winged vampire bats (*Diaemus youngi*) exchange contact calls. *Can. J. Zool.* **87**, 604–608. (doi:10.1139/Z09-051)

93. Leong KM, Ortolani A, Burks KD, Mellen JD, Savage A. 2003 Quantifying acoustic and temporal characteristics of vocalizations for a group of captive African elephants loxodonta Africana. *Bioacoustics* **13**, 213–231. (doi:10.1080/09524622.2003.9753499)

94. Soltis J, Leong K, Savage A. 2005 African elephant vocal communication I: antiphonal calling behaviour among affiliated females. *Anim. Behav.* **70**, 579–587. (doi:10.1016/j.anbehav.2004.11.015)

95. Yosida S, Okanoya K. 2009 Naked mole-rat is sensitive to social hierarchy encoded in antiphonal vocalization. *Ethology* **115**, 823–831. (doi:10.1111/j.1439-0310.2009.01677.x)

96. Buck J, Case JF. 1986 Flash control and female dialog repertoire in the firefly Photinus greeni. *Biol. Bull.* **170**, 176–197. (doi:10.2307/1541802)

97. Ewing AW. 1989 *Arthropod bioacoustics: neurobiology and behaviour*, p. 260. Edinburgh, UK: Edinburgh University Press.

98. Farabaugh SM. 1982 The ecological and social significance of duetting. In *Acoustic communication in birds* (eds DE Kroodsma, EH Miller), pp. 85–123. New York: NY: Academic Press.

99. Zimmermann U, Rheinlaender J, Robinson D. 1989 Cues for male phonotaxis in the duetting bushcricket Leptophyes punctatissima. *J. Comp.*

Physiol. A **164**, 621–628. (doi:10.1007/BF00614504)

100. Stumpner A, Meyer S. 2001 Songs and the function of song elements in four duetting bushcricket species (Ensifera, Phaneropteridae, Barbitistes). *J. Insect Behav.* **14**, 511–534. (doi:10.1023/A:1011176106943)

101. Forester DC, Harrison WK. 1987 The significance of antiphonal vocalisation by the spring peeper, pseudacris crucifer (*Amphibia, Anura*). *Behaviour* **103**, 1–15. (doi:10.1163/156853987X00233)

102. Grafe TU. 1996 The function of call alternation in the African reed frog (*Hyperolius marmoratus*): precise call timing prevents auditory masking. *Behav. Ecol. Sociobiol.* **38**, 149–158. (doi:10.1007/s002650050227)

103. Grafe TU. 1999 A function of synchronous chorusing and a novel female preference shift in an anuran. *Proc. R. Soc. Lond. B* **266**, 2331–2336. (doi:10.1098/rspb.1999.0927)

104. Wells KD. 1977 The social behaviour of anuran amphibians. *Anim. Behav.* **25**, 666–693. (doi:10.1016/0003-3472(77)90118-X)

105. Arak A. 1983 Vocal interactions, call matching and territoriality in a Sri Lankan treefrog, *Philautus leucorhinus* (Rhacophoridae). *Anim. Behav.* **31**, 292–302. (doi:10.1016/S0003-3472(83)80199-7)

106. Tobias ML, Viswanathan SS, Kelley DB. 1998 Rapping, a female receptive call, initiates male–female duets in the South African clawed frog. *Proc. Natl Acad. Sci. USA* **95**, 1870–1875. (doi:10.1073/pnas.95.4.1870)

107. Greenfield MD, Tourtellot MK, Snedden WA. 1997 Precedence effects and the evolution of chorusing. *Proc. R. Soc. Lond. B* **264**, 1355–1361. (doi:10.1098/rspb.1997.0188)

108. Wilson M, Wilson TP. 2005 An oscillator model of the timing of turn-taking. *Psychon. Bull. Rev.* **12**, 957–968. (doi:10.3758/BF03206432)

109. Lemasson A, Guilloux M, Rizaldi BS, Lacroix A, Koda H. 2013 Age- and sex-dependent contact call usage in Japanese macaques. *Primates* **54**, 283–291. (doi:10.1007/s10329-013-0347-5)

110. Gustison ML, Bergman TJ. 2016 Vocal complexity influences female responses to gelada male calls. *Sci. Rep.* **6**, 9. (doi:10.1038/srep19680)

111. Bergman TJ, Beehner JC, Cheney DL, Seyfarth RM. 2003 Hierarchical classification by rank and kinship in baboons. *Science* **302**, 1234–1236. (doi:10.1126/science.1087513)

112. Bates E, Benigni L, Bretherton I, Camaioni L, Volterra V. 1979 *The emergence of symbols: cognition and communication in infancy*, p. 387. New York, NY: Academic Press.

113. Bruner, J.S. 1981 Intention in the structure of action and interaction. In *Advances in infancy research* (ed. LP Lipsitt), pp. 41–56. NJ: Ablex, Norwood.

114. Schegloff EA. 2007 *Sequence organization in interaction, volume 1: a primer in conversation analysis*, p. 300. Cambridge, UK: Cambridge University Press.

115. Levinson SC. 2013 Recursion in pragmatics. *Language* **89**, 149–162. (doi:10.1353/lan.2013.0005)

116. Wilkinson R, Leudar I, Pika S. 2012 Requesting behaviours within episodes of active sharing. A new look on chimpanzee signalling. In *Developments in primate gesture research* (eds S Pika, K Liebal), pp. 199–221. Amsterdam, The Netherlands: John Benjamins Publishing Company.

117. Güntürkün O, Bugnyar T. 2016 Cognition without cortex. *Trends Cogn. Sci.* **20**, 291–303. (doi:10.1016/j.tics.2016.02.001)

118. Pika S, Bugnyar T. 2011 The use of referential gestures in ravens (*Corvus corax*) in the wild. *Nat. Commun.* **2**, 1–5. (doi:10.1038/ncomms1567)

119. Pepperberg IM. 1999 *The alex studies, cognitive and communicative abilities of grey parrots*, p. 448. Cambridge, MA: Harvard University Press.

120. Emery NJ, Clayton NS. 2004 The mentality of crows: convergent evolution of intelligence in corvids and apes. *Science* **306**, 1903–1907. (doi:10.1126/science.1098410)

121. Jarvis ED *et al.* 2005 Avian brains and a new understanding of vertebrate evolution. *Nat. Rev. Neurosci.* **6**, 151–159. (doi:10.1038/nrn1606)