



ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Special Issue: *Speech Rhythm in Ontogenetic, Phylogenetic, and Glossogenetic Development*

COMMENTARY

Interactive rhythms across species: the evolutionary biology of animal chorusing and turn-taking

Andrea Ravignani,^{1,2,3} Laura Verga,⁴ and Michael D. Greenfield^{5,6}

¹Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium. ²Institute for Advanced Study, University of Amsterdam, Amsterdam, the Netherlands. ³Research Department, Sealcentre Pieterburen, Pieterburen, the Netherlands.

⁴Faculty of Psychology and Neuroscience, Department NP&PP, Maastricht University, Maastricht, the Netherlands.

⁵Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas. ⁶Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI, CNRS UMR 9197, Université de Lyon/Saint-Etienne, Saint Etienne, France

Addresses for correspondence: Andrea Ravignani, Artificial Intelligence Lab, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium. andrea.ravignani@gmail.com; Michael D. Greenfield, Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045. michael.greenfield@univ-tours.fr

The study of human language is progressively moving toward comparative and interactive frameworks, extending the concept of *turn-taking* to animal communication. While such an endeavor will help us understand the interactive origins of language, any theoretical account for cross-species turn-taking should consider three key points. First, animal turn-taking must incorporate biological studies on animal *chorusing*, namely how different species coordinate their signals over time. Second, while concepts employed in human communication and turn-taking, such as intentionality, are still debated in animal behavior, lower level mechanisms with clear neurobiological bases can explain much of animal interactive behavior. Third, social behavior, interactivity, and cooperation can be orthogonal, and the alternation of animal signals need not be cooperative. Considering turn-taking a subset of chorusing in the rhythmic dimension may avoid overinterpretation and enhance the comparability of future empirical work.

Keywords: speech rhythm; interaction; cooperation; synchrony; language evolution; bioacoustics

Introduction

Interactivity is key in human and animal communication. However, the study of speech and language has traditionally focused on individual linguistic thought, production, and perception, rather than interactive communication.¹ This historical focus on individual language capacities, partly due to a generativist tradition in language sciences,^{2–4} is rapidly changing.^{5,6} Unanswered questions regarding how individual communicative rhythms are shaped by group interactions and modulated by social behavior are starting to be addressed.^{7–10} The study of conversation and *turn-taking* is taking center stage.^{11–13}

To understand the origins and evolution of language, the comparative approach targets analogous or homologous language-relevant traits in nonhuman animals.¹ Analogous evolutionary traits across species are those that stem from similar pres-

ures but have different origins,¹⁴ for instance, the presence of wings in birds and bats, or the capacity of both humans and bees to share referential information.¹ Instead, homologous evolutionary traits consist of evolutionarily old innovations in a common ancestor, for instance, the hind limbs of all land mammals used for movement, or their respiratory tract used for phonation.¹⁴ Within a cross-species framework, complex traits, such as human speech and language, are likely to be a mix of both: evolutionary old traits that can be found in many of our closest relatives and convergently evolved traits that only appear in clades distant from humans.^{1,15,16} In human communication, this comparative approach has been extended to turn-taking in recent years in an attempt to find how this trait has been shaped by evolution. As shown in Figure 1, this term is increasingly employed in the literature and may soon overtake *chorusing* by

doi: 10.1111/nyas.14230

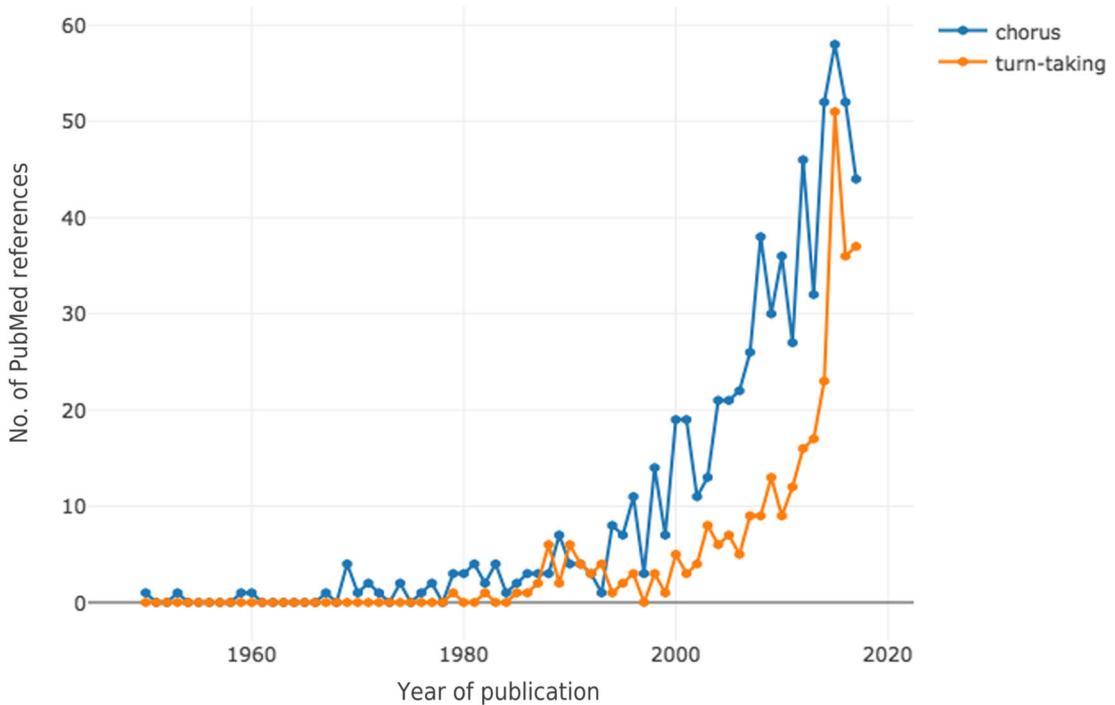


Figure 1. Number of PubMed entries for “chorus*” and “turn-taking” by year. The graph shows the development of chorusing and turn-taking research over the years, based on papers indexed by PubMed and published from 1950 to 2017. Visual inspection of the trends suggests a rise in publications related to turn-taking, which may soon surpass in volume those related to chorusing. Data were extracted from <http://dan.corlan.net/medline-trend.html>, and the graph was produced using the R package `plot_ly`. The regular expression “chorus*” was used in the search to account for the U.S. versus U.K. spelling differences.

number of publications. Among other aspects, research on turn-taking investigates the rhythm and timing of human and, more recently, animal rhythmic interactions.^{9,13,17–20} Across species, definitions of, and assumptions underlying, turn-taking strongly vary from simple call–response behavior to complex communicative patterns assuming “higher” cognition (Table 1).

These comparative efforts may offer valuable insights, but several issues need to be carefully monitored when investigating turn-taking in nonhuman species. In particular, we argue that (1) turn-taking may add a new (semantic and pragmatic) dimension to animal interactive communication, but in its temporal-rhythmic dimension it is nonetheless only a part of the general collective activity known as chorusing;^{19,26,27} (2) some concepts from human turn-taking are challenging to extend to all animals because they are still debated in (comparative) cognitive psychology, such as intentionality²⁸

or cognitive flexibility/complexity;²⁹ instead, attention-like processes, such as salience filters and competitive selection,³⁰ signal masking, and fixed time lags may suffice to explain much of animal interactive behavior; and (3) turn-taking—if defined as signal alternations, in all their nuances (Table 1)—does not require cooperation in living species or in hominin history. In fact, noncooperative coordination and competition often drive animal interactive displays.³¹ We illustrate these points by discussing and building upon some recent experiments and reviews of turn-taking in nonhuman animals.^{9,13,18,19,21,23,24} Our intention, we stress, is not to criticize a related field or a particular paper. Instead, we wish to spawn a new approach for examining *rhythm interaction across species*, an approach that will incorporate ideas from linguistics and cognition, via turn-taking, and from work in animal rhythmic communication, via chorusing.

Table 1. Different usages, definitions, or contextualization of turn-taking in the comparative literature

Species	Concepts (quotes)	Implications for defining turn-taking
Alston singing mice	<ul style="list-style-type: none"> • “Laboratory mice [...] fail to exhibit robust turn-taking behavior”²¹ “In contrast, marmoset pairs call antiphonally”²¹ • “In Alston’s singing mouse (<i>Scotinomys teguina</i>), we find a robust and rapid countersinging (~500 ms) that resembles the subsecond latencies of [...] the timing of vocal turn-taking evident in human conversation”²¹ • “Singing mice actively track conversations”²² • “Avoid interrupting each other”²² • “Rapidly perceiving song and precisely orchestrating a vocal reply”²² 	Subsecond latencies, overlap avoidance, antiphonal alternation, and robustness
Chimpanzees and bonobos	<ul style="list-style-type: none"> • “Human language is a fundamentally cooperative enterprise”²³ • “It has been suggested that it evolved as part of a larger adaptation of humans’ species unique forms of cooperation”²³ • “Turn-taking interactional practices embodying gestures to cooperatively share interest in an activity”²³ • “Structure and cognitive complexity underlying turn-taking sequences”²³ 	Cooperation, cognitive flexibility, intentionality, and cognitive complexity
Common marmosets	<ul style="list-style-type: none"> • “Never interrupt each other’s vocalizations during vocal exchanges [...] and will coordinate the timing of their calls relative to conspecifics”²⁴ • “Individuals will cease vocal interactions if a conspecific’s response latency is outside a particular period of time”²⁴ • “Turn-taking occurs within a few hundred milliseconds in human conversations.”²⁴ In marmosets: “a time scale of several seconds [...] limited to a single call type”²⁴ • “Vocal turn-taking—the repeated exchanges of vocalizations between any two individuals for an extended period of time (that is, not simply a call-and-response behaviour among mates or competitors)”²⁵ • “This turn-taking behaviour [...] exhibits features of a coupled oscillator system”²⁵ 	Overlap avoidance, limited time window to respond, beyond one call and response, and coupled-oscillator features
Cross-species	<ul style="list-style-type: none"> • “Alternating signal transmission between participants, with defined reply latency”¹⁸ • “Turn-taking behavior in animals can be classified into three categories based on the relationships between the signaler and the receiver: chorus, duet, or antiphony. Chorus involves males only; duet involves male-female pairs; and antiphony occurs in any animal combination.”¹⁸ • “Thus, one animal must send its signal after a preceding signal had ended. Sometimes the second signal is a response to the first. This alternation of signaler and receiver is referred to as turn-taking”¹⁸ 	Alternation, limited latency, chorus, duet or antiphonal behavior, and nonoverlapping response to a previous signal
Japanese macaques	<ul style="list-style-type: none"> • “Turn-taking, in which participants alternately reply to each other’s utterances”¹⁹ • “To enable this rapid turn-taking, the speaker must anticipate the timing of the partner”¹⁹ • “The male advertisement signals (flash or sound) of many insects and frogs have species-specific patterns, and result in synchrony or alternation (i.e., turn-taking, Greenfield <i>et al.</i> 1997; Lewis and Cratsley 2008)”¹⁹ • “As the term turn-taking refers to the exchange of communicative signals, such simultaneous signal production in nonhuman animals is also studied in the theoretical framework of ‘chorusing’”¹⁹ 	Response to a previous signal, anticipate the timing of a partner, potential synonym of chorusing, synchrony, and alternation
Male white-handed gibbons	<ul style="list-style-type: none"> • “The presence of antiphonal duetting in a species is not sufficient, however, to infer that the species possesses flexible turn-taking”²⁶ • “Unless it can be shown that duetting animals adjust their own calls or songs to a partner’s vocal output, then what seems to be turn-taking might instead be the chance result of simultaneous yet independent vocal production”²⁶ 	Beyond antiphonal duetting, temporal adjustment to a partner, and lack of co-occurrence of independent calls

Continued

Table 1. *Continued*

Species	Concepts (quotes)	Implications for defining turn-taking
Meerkats	<ul style="list-style-type: none"> • “Coordination can manifest as synchronization [...] but can also take the form of anti-synchronization or turn-taking”⁹ • “Examining the mechanism of turn-taking in animal groups [...] to determine whether the maintenance of multi-participant turn-taking can be a result of a spontaneous and cognitively simple process of self-assembly”⁹ • “Call inhibition over short timescales, which prevents mutual interference, and call excitation over longer timescales, which stimulates further group calling. These simple rules suggest that hierarchy formation and turn allocation are not required for achieving group-wide coordination of communication”⁹ 	Different from synchronization, a by-product of self-assembly and call inhibition or call excitation, turn allocation not needed

Chorusing versus turn-taking: what is the difference?

Turn-taking can be defined as “the orderly exchange of purely communicative signals or behaviors (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”¹³ (see also Tables 1 and 2). Animal chorusing can be defined as “Acoustic signalling produced collectively by a group of individuals whose activity is clustered in both space and time; choruses may be temporally structured in alternating or synchronous formats”³² (see also Table 2). It is incorrect to define chorusing as “a cacophony of sounds” or “synchronous production of the same call type by more than two individuals.”¹³ Indeed, cacophony is the exception, not the rule, in choruses,^{27,31} and only when synchronous instead of, for example, alternating²⁷ will a chorus be called a “synchronous chorus.”³¹ Adopting a too restrictive definition of chorusing undermines its importance and relegates it to a peculiar quirk of some animal species, of little relevance to the study of turn-taking. By contrast, we argue that a more productive way forward would be to consider turn-taking and chorusing as highly overlapping theoretical frameworks and methodological approaches to investigate one common concept—interactive communication—with an emphasis on its temporal and rhythmic dimension (for detailed reviews on chorusing and turn-taking, see Refs. 13, 27, and 31).

But what does chorusing have to offer to turn-taking researchers? The empirical study of animal choruses has been ongoing for almost a century

across species and modalities,^{7,8,33–39} and their study partly overlaps with animal turn-taking, a more recent and growing (Fig. 1) topic of comparative research.^{9,18,19,24,25} Differences among frameworks exist, and each framework can provide some insights that the other cannot.¹⁴ For instance, research on turn-taking is more concerned with cognitive flexibility, while chorusing focuses on the precise timing of interactive signals (see Table 2). Pragmatically, we argue that a systematic comparison between these two approaches may highlight the strengths and weaknesses of each, thus allowing mutual progress (Table 2, see also Refs. 14 and 40). Such comparison would point toward gaps in one field (e.g., which instances of turn-taking are competitive versus cooperative?) that can be filled by knowledge in the other (e.g., by performing model selection among mathematical frameworks of competitive versus cooperative chorusing).^{31,40–42} Importantly, the concept of chorusing can be applied to nonacoustic signaling,⁷ where it has been employed since its early origins.^{33,43,44} This is particularly convenient because language is inherently multimodal, rather than limited to the acoustic channel.

Toward a species-inclusive framework

A comparison between human and nonhuman interactive communication with a focus on rhythm is timely.^{8,13,50,54–56} However, at this early stage of animal turn-taking research, one needs to be careful in deciding which human concepts can be applied across species. For instance, taking turns may be important because one party may not be able to attend to a second when it is itself signaling

Table 2. Different approaches to and foci of animal interactive communication: similarities and differences between chorusing and turn-taking, often dependent on assumptions of what can be measured (see also Ref. 14)

Feature type	Turn-taking	Chorusing
Assumptions on measurability	Neural and “higher” cognitive processes ^{11,13}	Behavior and neurobiological processes ^{27,31}
Modality	Multimodal, as language is multimodal ^{11,23}	Mostly unimodal; depending on the species, focus on audition, movement, and/or vision ²⁷
Flexibility	Cognitive flexibility ^{11,13}	Temporal flexibility ^{26,31,45}
Assumption	Intentionality and some other cognitive traits ^{13,23}	Few or no mentalistic assumptions, rather focus on reactive versus predictive, and endogenous versus exogenous rhythms ²⁷
Timing	Absolute time delay between the offset of a signal and the onset of the next ⁴⁶ (see also Refs. 12, 47, and 48)	Relative phase of individual signal onsets with respect to the signals of other individuals ^{31,37,45}
Function	Relative communicative (proximate) function of two adjacent turns ¹¹	Evolutionary function(s) of interactive timing ^{31,35}
Group	Coordination between two individuals ⁴⁶	Emerging coordination patterns across two or more individuals ^{8,26,49}
Roles	Individual roles affected by the particular communicative exchange ¹¹	Sex, age, and social roles (e.g., dominant female/male) ^{50,51}
Semantics	Focus on the interplay between timing (and/or prosodic) and semantic (and/or pragmatic) content of a turn ^{11,52}	Little or no focus on semantics ³¹
“Syntax”	Turn-taking events are orderly organized, with emphasis on the variety of potential combinations, for example, ABC ≠ CAB ¹¹	Structural rules can govern the timing of events and “turns,” rather than their combination. When analyzed, the combination of turns are considered modulo their subcomponents, for example, (ABC)(ABC)AB = C(ABC)(ABC)A ^{37,45}
Mechanisms	Main focus on behavioral and neurocognitive mechanisms ¹¹	Focus on different explanatory levels, from genetics ^{10,53} to neurobiology and behavior ²⁷
Main methods	Discourse analysis of corpora of recorded interactions and neuroimaging ¹¹	Acoustic analyses of recordings, playback experiments, and so on ^{35,37,49}

(or lacks a neural forward model^{57,58}). In humans, sensory predictions of one’s own motor actions are used to attenuate the sensory effects of a movement (i.e., forward model)⁵⁷ so to better perceive external stimuli. For example, in a conversation, the attenuation of the sensory feedback due to one’s own voice may facilitate the perception of the interlocutor’s voice.⁵⁸ However, in other species, an individual may not be able to listen to another individual while signaling. This problem is particularly important in the acoustic modality because of physical masking, that is, when a stimulus hides another one with similar physical properties. Brains of humans and many other vertebrates generate forward models and efferent copies to solve this issue,^{58,59} but it is unclear whether this system may be sufficiently developed in other taxa such as arthropods.⁶⁰

Hence, it is currently unknown whether this concept may be broadly applicable to a turn-taking comparative framework.⁴

A potential risk in applying turn-taking frameworks to animal communication may be to neglect these important, albeit “lower level,” issues to focus more on cognitive, mentalistic frameworks.⁶¹ Concepts like intentionality, cognitive flexibility, and

^aWhile less critical, the issue of masking also exists in visual signaling.¹² An intriguing research direction could be a cross-species comparison of signal masking between auditory and visual modalities.^{7,12} Such research would test the hypotheses of phylogenetic effects of efferent copies or modality specificity on successful interactive communication.

cognitive complexity are sometimes used in animal turn-taking research.^{13,23} Especially in apes, these concepts seem quite suitable.²³ However, if the aim is a broad comparative approach, are these appropriate yardsticks for comparison? For instance, what are the requirements for a cricket's communication act to be classified as intentional? The field of comparative cognition and animal behavior has been debating for decades which species show intentional behavior.^{62–66} some argue that only humans possess intentionality, while others see no reason to exclude any animal, or even bacteria, species,²⁸ and robots.⁶⁷ Likewise, without falling back into pure behaviorism,⁶¹ it is still unclear which species the cognitivist framework should include,²⁹ how “cognitive complexity” should be measured,⁶⁸ and by whom.⁶⁹ To ensure the cross-species study of turn-taking rests on solid ground, only species-inclusive frameworks, ideally extendable to all animals, should be initially adopted. For instance, a framework applicable to all signaling species is the use of timing in social group behavior.^{9,70,71} Nonetheless, the species-inclusive framework we advocate may obscure fine-grained differences between taxa, and even within a taxon subject to diverse environmental and social conditions. Hence, once this global data set is developed, focused analyses of specific sections could allow us to avoid anthropocentrism while seeking to understand where cooperation via shared goals is a factor in signal alternation.

Turn-taking does not imply cooperation

Many studies and reviews of human language conjecture its origin in cooperation.^{23,72,73} These conjectures may account for the current trend linking behavioral alternation across a wide range of nonhuman animals to cooperation and referring to such alternation as turn-taking.^{13,23} While we are sympathetic toward *interactive* accounts of the evolution of language and speech,^{6,14,27,74} turn-taking, cooperation, and language origins have not been shown to be so strongly connected as depicted in some turn-taking work.^{13,23} Across nonhuman animals, acoustic turn-taking is certainly widespread, but it is not necessarily cooperative.^{39,41,45} Cooperation may be an important element in some aspects of avian duetting,⁷⁵ primate communication,^{13,23} and in insect male–female exchanges. However, many examples of male–male interchanges involve competition.^{35,76–80} In insects and anurans, what

appears to be turn-taking between males has been experimentally shown to be a way to clearly assess a rival neighbor, and ensure that a rival neighbor hears you correctly and takes appropriate action—which may be leaving and avoiding an attack.⁸¹ Thus, when prolonged call alternation occurs, the matched calls reflect individuals who have evaluated one another until that juncture as rather equivalent in physical prowess or motivation.⁸² When one of the individuals eventually does not respond, he is likely to have tired and will shortly leave, thereby avoiding the high probability of being attacked by a stronger neighbor. In this context, the call matching during the alternation phase of the interaction is quite unlike turn-taking in typical human conversations.^b Cooperation would demand that a faster individual slow down to accommodate the slower one, an adjustment common in humans^{86,87} and experimentally achievable in other mammal species (e.g., sea lions⁸⁸), which does not seem to occur in anuran and insect species.⁸⁹ Another case where turn-taking may at first appear as a cooperative venture but is really a by-product of competition occurs where leading calls are preferred by receivers (precedence effect). Under such circumstances, each caller in a local group actively avoids being a follower. In many animal species, the avoidance entails a “phase-delay” mechanism wherein a caller inhibits his central rhythm generator when he hears a neighbor's call.^{37,90} The caller then rebounds from inhibition at the end of neighbor's call and thereby calls before the neighbor's next call.^{37,90} When two individuals call with similar rhythms and phase-delay adjustments, alternation—arguably analogous to turn-taking^{3,27}—emerges by default. From an evolutionary perspective, phenomena that appear to be cooperative today may have their origins in phenomena that most decidedly were not. For example, mutualism between species may have begun as parasitism.⁹¹ Likewise, animal alarm calls, which provide direct benefit to the receiver but not the caller, might have originated to selfishly manipulate conspecifics' behavior.^{92,93}

In brief, alternative hypotheses—both competitive and cooperative—may account for turn-taking

^bHuman language is flexible, however, and in some cases confrontational or “competitive” turn-taking does occur.^{83–85}

behavior.^{40,41} Likewise, different individual behaviors can be mapped to a range of interactivity patterns.^{40,74,90} This mapping is complex, non-linear, and a potential subject for exciting future research. It must include at least four explanatory levels, illustrating (1) how neural processes occur within an individual nervous system (e.g., oscillations, predictive coding, and climbing neural activity^{94–98}); (2) possibly supporting different cognitive mechanisms;^{13,14,30,40,60,74} (3) which result in a range of individual timing patterns that can be measured behaviorally;^{12,37,45,46,49} and that (4) shape and are shaped by interindividual interaction.^{14,39,71,98}

Conclusions

To conclude, we welcome and support a cross-species approach to understanding turn-taking. However, a clear common definition of turn-taking should be employed across studies (see Table 1). We agree with our colleagues who are extending a human turn-taking framework to other species that testable frameworks are important, that there is a shortage of data, and that, at the current stage, it cannot be decided which species show elements of human conversational turn-taking.^{13,18} We argue that findings do not need to be weighted by phylogenetic distance from humans, and many examples of animal turn-taking^{9,19,21,24} are still interesting in their own merit without advocating cooperation or social cognition. “Lower level” processes, such as attention,³⁰ signal masking,^{35,77} and fixed time lags³¹ may suffice to explain much of animal turn-taking behavior. In particular, animal chorusing, as long as this is not defined as “a cacophony of sounds,”¹³ may provide a broader biological foundation for turn-taking without assuming intentionality. The intention may very well be present in some instances of turn-taking and chorusing, but there are many cases where it is likely not. As a new hybrid field is born, its cross-species approach to turn-taking and chorusing will feature a real interdisciplinary blend among biology, cognition, and neuroscience.

Acknowledgments

We are grateful to Katie Mudd, Koen de Reus, and Ruth Sonnweber for insightful comments on this paper. A.R. has received funding from the European Union’s Horizon 2020 research and innovation pro-

gramme under the Marie Skłodowska-Curie grant agreement no 665501 with the research Foundation Flanders (FWO) (Pegasus² Marie Curie fellowship 12N5517N).

Author contributions

A.R., L.V., and M.D.G. discussed and wrote the manuscript. All authors gave final approval for publication.

Competing interests

The authors declare no competing interests.

References

1. Fitch, W.T. 2010. *The Evolution of Language*. Cambridge University Press.
2. Berwick, R.C. & N. Chomsky. 2017. Why only us: recent questions and answers. *J. Neurolinguist.* **43**: 166–177.
3. Ravignani, A., B. Thompson & P. Filippi. 2018. The evolution of musicality: what can be learned from language evolution research? *Front. Neurosci.* **12**: 20.
4. Chomsky, N. 2017. The language capacity: architecture and evolution. *Psychonom. Bull. Rev.* **24**: 200–203.
5. Holler, J. & S.C. Levinson. 2019. Multimodal language processing in human communication. *Trends Cogn. Sci.* **23**: 639–652.
6. Verga, L. & S.A. Kotz. 2019. Putting language back into ecological communication contexts. *Lang. Cogn. Neurosci.* **34**: 536–544.
7. Backwell, P.R.Y. 2018. Synchronous waving in fiddler crabs: a review. *Curr. Zool.* **65**: 83–88.
8. De Gregorio, C., A. Zanoli, D. Valente, *et al.* 2018. Female indris determine the rhythmic structure of the song and sustain a higher cost when the chorus size increases. *Curr. Zool.* **65**: 89–97.
9. Demartsev, V., A. Strandburg-Peshkin, M. Ruffner, *et al.* 2018. Vocal turn-taking in meerkat group calling sessions. *Curr. Biol.* **28**: 3661–3666.e3663.
10. Torti, V., G. Bonadonna, C. De Gregorio, *et al.* 2017. An intra-population analysis of the indris’ song dissimilarity in the light of genetic distance. *Sci. Rep.* **7**: 10140.
11. Holler, J., K.H. Kendrick, M. Casillas, *et al.* 2015. Turn-taking in human communicative interaction. *Front. Psychol.* **6**: 1919.
12. de Vos, C., F. Torreira & S.C. Levinson. 2015. Turn-timing in signed conversations. Coordinating stroke-to-stroke turn boundaries. *Front. Psychol.* **6**. <https://doi.org/10.3389/fpsyg.2015.00268>.
13. Pika, S., R. Wilkinson, K.H. Kendrick, *et al.* 2018. Taking turns: bridging the gap between human and animal communication. *Proc. Biol. Sci.* **285**. <https://doi.org/10.1098/rspb.2018.0598>.
14. Kotz, S., A. Ravignani & W.T. Fitch. 2018. The evolution of rhythm processing. *Trends Cogn. Sci.* **22**: 896–910.

15. Fitch, W.T., B. de Boer, N. Mathur, *et al.* 2016. Monkey vocal tracts are speech-ready. *Sci. Adv.* **2**: e1600723.
16. Seed, A., N. Emery & N. Clayton. 2009. Intelligence in corvids and apes: a case of convergent evolution? *Ethology* **115**: 401–420.
17. Takahashi, D.Y., D.Z. Narayanan & A.A. Ghazanfar. 2013. Coupled oscillator dynamics of vocal turn-taking in monkeys. *Curr. Biol.* **23**: 2162–2168.
18. Yoshida, S. & K. Okanoya. 2005. Evolution of turn-taking: a bio-cognitive perspective. *Cognit. Stud.* **12**: 153–165.
19. Katsu, N., K. Yamada, K. Okanoya, *et al.* 2019. Temporal adjustment of short calls according to a partner during vocal turn-taking in Japanese macaques. *Curr. Zool.* **65**: 99–105.
20. Dohmen, D. & S.R. Hage. 2019. Limited capabilities for condition-dependent modulation of vocal turn-taking behavior in marmoset monkeys. *Behav. Neurosci.* **133**: 320.
21. Okobi, D.E., A. Banerjee, A.M. Matheson, *et al.* 2019. Motor cortical control of vocal interaction in neotropical singing mice. *Science* **363**: 983–988.
22. Vahaba, D.M. 2019. Singing mice take turns talking. *J. Exp. Biol.* **222**: jeb192765.
23. Fröhlich, M., P. Kuchenbuch, G. Müller, *et al.* 2016. Unpeeling the layers of language: bonobos and chimpanzees engage in cooperative turn-taking sequences. *Sci. Rep.* **6**: 25887.
24. Chow, C.P., J.F. Mitchell & C.T. Miller. 2015. Vocal turn-taking in a non-human primate is learned during ontogeny. *Proc. Biol. Sci.* **282**: 20150069.
25. Takahashi, D.Y., A.R. Fenley & A.A. Ghazanfar. 2016. Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philos. Trans. R. Soc. B* **371**: 20150370.
26. Terleph, T.A., S. Malaivijitnond & U.H. Reichard. 2018. Male white-handed gibbons flexibly time duet contributions. *Behav. Ecol. Sociobiol.* **72**: 16.
27. Ravignani, A., D.L. Bowling & W.T. Fitch. 2014. Chorusing, synchrony and the evolutionary functions of rhythm. *Front. Psychol.* **5**: 1118.
28. Fitch, W.T. 2008. Nano-intentionality: a defense of intrinsic intentionality. *Biol. Philos.* **23**: 157–177.
29. Webb, B. 2012. Cognition in insects. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**: 2715–2722.
30. Nityananda, V. 2016. Attention-like processes in insects. *Proc. Biol. Sci.* **283**: 20161986.
31. Hartbauer, M. & H. Römer. 2016. Rhythm generation and rhythm perception in insects: the evolution of synchronous choruses. *Front. Neurosci.* **10**. <https://doi.org/10.3389/fnins.2016.00223>.
32. Greenfield, M.D. 2002. *Signalers and Receivers: Mechanisms and Evolution of Arthropod Communication*. Oxford University Press.
33. Buck, J. 1938. Synchronous rhythmic flashing in fireflies. *Q. Rev. Biol.* **13**: 301–314.
34. Winfree, A.T. 1986. *Timing of Biological Clocks*. Henry Holt and Company.
35. Grafe, T.U. 1999. A function of synchronous chorusing and a novel female preference shift in an anuran. *Proc. R. Soc. Lond. B Biol. Sci.* **266**: 2331–2336.
36. Strogatz, S.H. 2003. *Sync: The Emerging Science of Spontaneous Order*. Hyperion.
37. Greenfield, M.D. & I. Roizen. 1993. Katydid synchronous chorusing is an evolutionarily stable outcome of female choice. *Nature* **364**: 618–620.
38. Aihara, I., H. Kitahata, K. Yoshikawa, *et al.* 2008. Mathematical modeling of frogs' calling behavior and its possible application to artificial life and robotics. *Artif. Life Robot.* **12**: 29–32.
39. Ota, K., I. Aihara & T. Aoyagi. 2019. Interaction mechanisms quantified from dynamical features of frog choruses. arXiv preprint arXiv:1907.11403.
40. Ravignani, A. & K. de Reus. 2019. Models of animal rhythmic signalling. *Evol. Bioinform.* **15**: 1–14.
41. Greenfield, M.D. 1994. Cooperation and conflict in the evolution of signal interactions. *Annu. Rev. Ecol. Syst.* **25**: 97–126.
42. Hartbauer, M., S. Kratzer, K. Steiner & H. Römer. 2005. Mechanisms for synchrony and alternation in song interactions of the bushcricket *Mecopoda elongata* (Tettigoniidae: Orthoptera). *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **191**: 175–188.
43. Buck, J. 1988. Synchronous rhythmic flashing in fireflies. II. *Q. Rev. Biol.* **63**: 265–287.
44. Buck, J. & E. Buck. 1968. Mechanism of rhythmic synchronous flashing of fireflies: fireflies of Southeast Asia may use anticipatory time-measuring in synchronizing their flashing. *Science* **159**: 1319–1327.
45. Ravignani, A. 2019. Timing of antisynchronous calling: a case study in a harbor seal pup (*Phoca vitulina*). *J. Comp. Psychol.* **133**: 272–277.
46. Stivers, T., N.J. Enfield, P. Brown, *et al.* 2009. Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci. USA* **106**: 10587–10592.
47. Levinson, S.C. & F. Torreira. 2015. Timing in turn-taking and its implications for processing models of language. *Front. Psychol.* **6**: 731.
48. Hilbrink, E.E., M. Gattis & S.C. Levinson. 2015. Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Front. Psychol.* **6**: 1492.
49. Aihara, I., T. Mizumoto, T. Otsuka, *et al.* 2014. Spatio-temporal dynamics in collective frog choruses examined by mathematical modeling and field observations. *Sci. Rep.* **4**: 3891.
50. Gamba, M., V. Torti, V. Estienne, *et al.* 2016. The indris have got rhythm! Timing and pitch variation of a primate song examined between sexes and age classes. *Front. Neurosci.* **10**. <https://doi.org/10.3389/fnins.2016.00249>.
51. Clink, D.J., J.S. Tasirin & H. Klinck. 2019. Vocal individuality and rhythm in male and female duet contributions of Gursky's spectral tarsier. *Curr. Zool.* <https://doi.org/10.1093/cz/zoz035>.
52. Corps, R.E., A. Crossley, C. Gambi, *et al.* 2018. Early preparation during turn-taking: listeners use content predictions to determine what to say but not when to say it. *Cognition* **175**: 77–95.
53. Mahamoud-Issa, M., T. Marin-Cudraz, V. Party, *et al.* 2017. Phylogeographic structure without pre-mating barriers: do

- habitat fragmentation and low mobility preserve song and chorus diversity in a European bushcricket? *Evol. Ecol.* **31**: 865–884.
54. Ravignani, A., S. Dalla Bella, S. Falk, *et al.* 2019. Rhythm in speech and animal vocalizations: a cross-species perspective. *Ann. N.Y. Acad. Sci.* **1453**: 79–98.
 55. Filippi, P. 2016. Emotional and interactional prosody across animal communication systems: a comparative approach to the emergence of language. *Front. Psychol.* **7**: 1393.
 56. Moore, R.K. 2019. Vocal interactivity in crowds, flocks and swarms: implications for voice user interfaces. arXiv preprint arXiv:1907.11656.
 57. Blakemore, S.-J., D. Wolpert & C. Frith. 2000. Why can't you tickle yourself? *Neuroreport* **11**: R11.
 58. Greenlee, J.D., A.W. Jackson, F. Chen, *et al.* 2011. Human auditory cortical activation during self-vocalization. *PLoS One* **6**: e14744.
 59. Eliades, S.J. & X. Wang. 2003. Sensory–motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* **89**: 2194–2207.
 60. Webb, B. 2004. Neural mechanisms for prediction: do insects have forward models? *Trends Neurosci.* **27**: 278–282.
 61. Shettleworth, S.J. 2001. Animal cognition and animal behavior. *Anim. Behav.* **61**: 277–286.
 62. Allen, C. & M. Bekoff. 1995. Cognitive ethology and the intentionality of animal behavior. *Mind Lang.* **10**: 313–328.
 63. Heyes, C. & A. Dickinson. 1990. The intentionality of animal action. *Mind Lang.* **5**: 87–103.
 64. Byrne, R.W. 1999. Imitation without intentionality. Using string parsing to copy the organization of behavior. *Anim. Cogn.* **2**: 63–72.
 65. Schel, A.M., S.W. Townsend, Z. Machanda, *et al.* 2013. Chimpanzee alarm call production meets key criteria for intentionality. *PLoS One* **8**: e76674.
 66. Searle, J. 1980. Intrinsic intentionality. *Behav. Brain Sci.* **3**: 450–457.
 67. Kopp, L. & P. Gärdenfors. 2001. *Attention as a Minimal Criterion of Intentionality in Robots*. Citeseer.
 68. Chittka, L., S.J. Rossiter, P. Skorupski & C. Fernando. 2012. What is comparable in comparative cognition? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**: 2677–2685.
 69. Kokko, H. 2017. Give one species the task to come up with a theory that spans them all: what good can come out of that? *Proc. Biol. Sci.* **284**: 20171652.
 70. Schirmer, A., W.H. Meck & T.B. Penney. 2016. The socio-temporal brain: connecting people in time. *Trends Cogn. Sci.* **20**: 760–772.
 71. Ravignani, A. & S. Kotz. 2017. Evolution of temporal interaction: a comparative approach to social timing. *PeerJ Preprints* **5**: e3275v3271.
 72. Ghazanfar, A.A. & D.Y. Takahashi. 2014. The evolution of speech: vision, rhythm, cooperation. *Trends Cogn. Sci.* **18**: 543–553.
 73. Smith, E.A. 2010. Communication and collective action: language and the evolution of human cooperation. *Evol. Hum. Behav.* **31**: 231–245.
 74. Ravignani, A., D. Bowling & S. Kirby. 2014. The psychology of biological clocks: a new framework for the evolution of rhythm. In *Presented at the Evolution of Language: Proceedings of the 10th International Conference*.
 75. Mann, N.I., K.A. Dingess & P. Slater. 2006. Antiphonal four-part synchronized chorusing in a neotropical wren. *Biol. Lett.* **2**: 1–4.
 76. Tobias, M.L., S.S. Viswanathan & D.B. Kelley. 1998. Rapping, a female receptive call, initiates male–female duets in the South African clawed frog. *Proc. Natl. Acad. Sci. USA* **95**: 1870–1875.
 77. Grafe, T.U. 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.
 78. Forester, D.C. & W.K. Harrison. 1987. The significance of antiphonal vocalisation by the spring peeper, *Pseudacris crucifer* (Amphibia, Anura). *Behavior* **103**: 1–15.
 79. Wells, K.D. 1977. The social behavior of anuran amphibians. *Anim. Behav.* **25**: 666–693.
 80. Arak, A. 1983. Vocal interactions, call matching and territoriality in a Sri Lankan treefrog, *Philautus leucorhinus* (Rhacophoridae). *Anim. Behav.* **31**: 292–302.
 81. Greenfield, M.D. & R.L. Minckley. 1993. Acoustic dueling in tarbush grasshoppers: settlement of territorial contests via alternation of reliable signals. *Ethology* **95**: 309–326.
 82. Greenfield, M.D. 2015. Signal interactions and interference in insect choruses: singing and listening in the social environment. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **201**: 143–154.
 83. Ulijn, J.M. & L. Xiangling. 1995. Is interrupting impolite? Some temporal aspects of turn-taking in Chinese–Western and other intercultural business encounters. *Text* **15**: 589–628.
 84. Byrnes, H. 1986. Interactional style in German and American conversations. *Text* **6**: 189–206.
 85. Hayashi, R. 1988. Simultaneous talk—from the perspective of floor management of English and Japanese speakers. *World Englishes* **7**: 269–288.
 86. Wing, A.M., S. Endo, A. Bradbury, *et al.* 2014. Optimal feedback correction in string quartet synchronization. *J. R. Soc. Interface* **11**: 20131125.
 87. Bowling, D.L., C.T. Herbst & W.T. Fitch. 2013. Social origins of rhythm? Synchrony and temporal regularity in human vocalization. *PLoS One* **8**: e80402.
 88. Rouse, A.A., P.F. Cook, E.W. Large & C. Reichmuth. 2016. Beat keeping in a sea lion as coupled oscillation: implications for comparative understanding of human rhythm. *Front. Neurosci.* **10**. <https://doi.org/10.3389/fnins.2016.00257>.
 89. Greenfield, M.D. 1994. Synchronous and alternating choruses in insects and anurans: common mechanisms and diverse functions. *Am. Zool.* **34**: 605–615.
 90. Greenfield, M.D., M.K. Tourtellot & W.A. Snedden. 1997. Precedence effects and the evolution of chorusing. *Proc. R. Soc. Lond. B Biol. Sci.* **264**: 1355–1361.
 91. Shapiro, J.W. & P.E. Turner. 2018. Evolution of mutualism from parasitism in experimental virus populations. *Evolution* **72**: 707–712.
 92. Charnov, E.L. & J.R. Krebs. 1975. The evolution of alarm calls: altruism or manipulation? *Am. Nat.* **109**: 107–112.

93. Smith, J.M. 1965. The evolution of alarm calls. *Am. Nat.* **99**: 59–63.
94. Wittmann, M. 2013. The inner sense of time: how the brain creates a representation of duration. *Nat. Rev. Neurosci.* **14**: 217.
95. Arnal, L.H., V. Wyart & A.-L. Giraud. 2011. Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat. Neurosci.* **14**: 797.
96. Rimmele, J.M., B. Morillon, D. Poeppel, *et al.* 2018. Proactive sensing of periodic and aperiodic auditory patterns. *Trends Cogn. Sci.* **22**: 870–882.
97. Giraud, A.-L. & D. Poeppel. 2012. Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* **15**: 511.
98. Wilson, M. & T.P. Wilson. 2005. An oscillator model of the timing of turn-taking. *Psychonom. Bull. Rev.* **12**: 957–968.