ANNUAL REVIEWS

Animal Communication in Linguistic and Cognitive Perspective

Thom Scott-Phillips[1] and Christophe Heintz[2]

[1]Institute for Logic, Cognition, Language, and Information, San Sebastian, Spain; email: thom.scottphillips@gmail.com

[2]Department of Cognitive Science, Central European University, Budapest, Hungary

ANNUAL REVIEWS CONNECT

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

animal communication, comparative cognition, Washburn, Ladygina-Kohts, Grice, duality of patterning, generativity

Abstract

Detailed comparative studies have revealed many surface similarities between linguistic communication and the communication of nonhumans. How should we interpret these discoveries in linguistic and cognitive perspective? We review the literature with a specific focus on analogy (similar features and function but not shared ancestry) and homology (shared ancestry). We conclude that combinatorial features of animal communication are analogous but not homologous to natural language. Homologies are found instead in cognitive capacities of attention manipulation, which are enriched in humans, making possible many distinctive forms of communication, including language use. We therefore present a new, graded taxonomy of means of attention manipulation, including a new class we call Ladyginian, which is related to but slightly broader than the more familiar class of Gricean interaction. Only in the latter do actors have the goal of revealing specifically informative intentions. Great ape interaction may be best characterized as Ladyginian but not Gricean.

1. INTRODUCTION

Languages: sets of communicative conventions for ostensive communication, (self-)organized in structured networks

Analogy and homology: Homologous features are similar in form because they derive from a common ancestral trait; analogous features are similar in form because they fulfill similar functions but have evolved independently

In recent decades the empirical study of animal communication from a linguistic perspective has advanced considerably, moving well beyond the classic examples featured in many linguistic textbooks, such as the honeybee waggle dance, vervet monkey alarm calls, and the ape language experiments of the mid-twentieth century (for summaries and perspectives on this earlier research, see, e.g., Seyfarth & Cheney 2003; Munz 2005; Anderson 2006; Radick 2007; Hurford 2007, 2012; Pepperberg 2017; Rendall 2021). Recent findings have revealed similarities with natural languages and natural language use at all levels of analysis, from phonology to pragmatics. Perhaps the most striking discovery has been the extent to which nonhuman communication systems have some of the same surface structures as natural languages. As one representative example, the calls of chestnut-crowned babblers (*Pomatostomus ruficeps*) comprise individual elements that are perceptibly distinct within calls, are perceptibly equivalent across calls, and convey no relevant information on their own; but when used in combination these elements form constructions that do convey relevant information (Engesser et al. 2015, 2019). This is the collection of qualities commonly known as duality of patterning, which is widely identified as a distinctive and foundational property of natural languages.[1] The discovery of the same collection of qualities in the calls of a songbird is one of several intriguing developments that raise deep and important questions of interpretation.

Yet the question of exactly how to accommodate such discoveries and future breakthroughs into linguistic theory remains unresolved and is often contentious. Some research programs use analytical tools developed within linguistics to study animal communication systems simply for their own sake, with no prior commitment to any particular interpretation about comparisons with natural languages (e.g., Schlenker et al. 2016). Other research programs seem more committed to assumptions that similarities between animal communication systems and natural languages reflect something important about the nature of language and languages themselves. In some cases these assumptions are tacit and unexamined; in others they are long-held theoretical commitments; but either way they influence interpretation. For instance, the data on chestnut-crowned babblers, summarized above, are approached and interpreted very differently by linguists working in the generative tradition than by the team of biologists and linguists who conducted the research (for debate, see, e.g., Huybregts 2020, Engesser & Fitch 2021). Another example is debate over the possible existence of reference in nonhuman primate communication (e.g., Wheeler & Fischer 2012, Scarantino & Clay 2015). Put simply, there is little consensus about how to interpret the findings of animal communication research from a linguistic perspective.

Here we review and reinterpret these literatures, with a specific focus on the distinction between homology and analogy. In biology, homologous features are similar in form because they derive from a common ancestral trait. Analogous features are similar in form because they fulfil similar functions, but have evolved independently. For instance, the bones in the front flipper of a whale are homologous to the bones in the human arm: Both derive from a common ancestral trait.

---

[1]Some further elaboration for nonlinguists: In spoken language, the elements are phonemes, or individual sounds, such as /k/ and /b/. Phonemes are perceptibly distinct within utterances (we can differentiate /k/ from /b/), are perceptibly equivalent across utterances (we can identify different specific instances of /k/ as tokens of the same type, and ditto for /b/), and convey no relevant information on their own (/k/ and /b/ are each meaningless). But when used in combination, these elements form constructions that do convey relevant information. /kju:b/ ('cube') is meaningful in English, even though /k/, /j/, /u:/, and /b/ are not; and this meaningful construction can in turn be combined with other meaningful constructions. This set of properties is commonly called duality of patterning, although researchers' use of the term varies (de Boer et al. 2012; for phonological organization in sign languages, see Sandler 2012).

WHAT IS LANGUAGE EVOLUTION?

In the most general sense, the field of language evolution is best characterized as the study of how and why the phenomenon of language (whatever that is) can come to be (Dediu & de Boer 2016). One way to elaborate this domain is by distinguishing the following three things (see also Scott-Phillips & Kirby 2010, Scott-Phillips 2017):

1. The biological evolution of whatever cognitive capacities are necessary to generate, acquire, and use languages.
2. The transition from very simple systems, with few if any of the characteristic features of languages (e.g., no or very little syntax), into a system that is characteristically linguistic.
3. The change from one characteristically linguistic system to another.

Historical linguistics is classically concerned with point 3. Language evolution is centrally concerned with points 1 and 2. That is—to use terms from elsewhere in this article—language evolution is concerned with the various shifts entailed by moves to Gricean, Lewisian, and ultimately Saussurian systems (see **Figure 1**, below). This is not to deny that the study of point 3 has findings and insights relevant to point 2—on the contrary, in fact.

In contrast, the wings of a butterfly and the wings of a bird are analogous but not homologous: They perform similar functions but do not derive from a common ancestral trait. The question is: Which forms of animal communication and interaction, if any, are homologous with the cognitive capacities that humans employ in ordinary language use, and which are analogous? This question is fundamental to the evolutionary, cognitive, and interactive origins of human communication, languages, and language use (see the sidebar titled What Is Language Evolution?).

We argue for two complementary conclusions. In Section 2 we consider combinatorics, focusing on key similarities and dissimilarities between natural language and the vocal communication of other species. We adopt this focus because the vocal domain is where many of the most striking similarities between natural language and nonhuman communication have been observed, such as in the example of chestnut-crowned babblers described above. We conclude that the surface similarities shared between natural language and the communication systems of other species are analogous but not homologous. That is, there are similarities in the structures of communicative systems but not in the cognitive capacities that underpin communicative systems. Thus, in Section 3 we consider meaning and attention manipulation. We describe a key puzzle of great ape gestural communication: How can we describe great ape gesture in a way that accounts for both its behavioral similarity with many human behaviors and its limited range and scope relative to the human case? We then suggest a resolution of this puzzle, with a novel analysis based on social cognition and, in particular, different modes of attention manipulation. We conclude that great apes manipulate others' attention in ways that are homologous with human capacities but have been enriched further in the *Homo* lineage, making humans "language ready."

In other words, we argue that evolutionary continuity exists between humans and other species in terms of cognitive capacities of attention manipulation but not in terms of communication systems themselves. The origins of human language can be found in distinctly human forms of attention manipulation, and the combinatorial richness of human language is an exploitation of this difference, not its source (see also Heintz & Scott-Phillips 2023). As such, the discovery of combinatorial communication systems in nonhumans is revealing not of any direct evolutionary continuity with syntax or phonology but rather of the many and varied features of combinatorial cognition that can shape natural languages as they emerge, change, and evolve (see Section 4). Related but nevertheless distinct topics that we do not have space to review include the comparative biology of speech and speech apparatus (for reviews, see, e.g., Ghazanfar & Rendall 2008, Fitch 2018), Gricean communication in human infancy (e.g., Tomasello et al. 2007, Bohn & Frank

Combinatorial communication: individual elements used together for a function that is not simply the sum of the functions of the individual elements

Gricean: refers to communication by means of overt intentionality (i.e., making informative intentions themselves overt)

2019), and artificial language learning experiments in nonhumans (e.g., Fitch & Friederici 2012, ten Cate 2017).

2. COMBINATORICS

2.1. Combinatorics Are Copious in Nonhumans

A minimal characterization of combinatorics is that individual elements, each of which has a particular function if produced alone, can be used together for a function that is not simply the sum of the functions of the individual elements (Scott-Phillips & Blythe 2013; on communication in general, see Scott-Phillips & Kirby 2013). Duality of patterning, for instance, is combinatoric on two distinct levels, classically described as phonology and morphology (de Boer et al. 2012). At one level, individual elements are combined into new structures, which are not simply the sum of the elements (morphemes are not simply the sum of their phonemes), and these combinations are themselves individual elements at a higher level (sentences are not simply the sum of their morphemes).

How can combinatorics be investigated in other species? The typical method is observation in natural environments followed by controlled playback experiments. In such experiments, natural stimuli are recorded and played to target animals under varying conditions, and the responses are then documented and measured (Radick 2007, Fischer 2017). To investigate combinatorics in particular, the recorded natural stimuli must be divided into parts and then systematically recombined to generate appropriate experimental stimuli. Such experiments have been conducted both in the wild and in more controlled settings, such as with monkeys and apes in captivity.

Here is a specific example to help illustrate the general method and hence elaborate the nature of the data on which conclusions are based (Arnold & Zuberbühler 2006, 2008). Male putty-nosed monkeys (*Cercopithecus nictitans*) produce two acoustically distinct loud calls (pyows and hacks), depending on external events. In particular, pyows are used primarily when a leopard is observed, and hacks primarily when crowned eagles are observed. Observation of natural behavior suggested that combinations of these two calls may have a different function. These naturally occurring combinations were recorded and artificially played back to monkeys, and their responses were measured and compared with control conditions of pyows only, hacks only, and no calls. This method was repeated for artificially created pyow–hack combinations. The results revealed that pyow–hack combinations elicited a wholly different behavioral response, namely group travel. This is evidence of combinatorial communication because the combination of calls achieves an effect that is different from the sum, or combination, of the effects of the component parts.

Using such methods, researchers have identified combinatorics of various kinds in the natural communication systems of many species (for detailed reviews, analyses, and debate, see, e.g., Bolhuis et al. 2010, Berwick et al. 2011, Hurford 2012, Schlenker et al. 2016, Engesser & Townsend 2019, Heesen et al. 2019, Suzuki & Zuberbühler 2019, Leroux & Townsend 2020, Zuberbühler 2020). Much research has focused on nonhuman primates and birdsong, but combinatorics have been discovered in a very wide range of species, including, for instance, whales and bacteria. Most discoveries have occurred in the past 15 years, and given the present rate of discovery it is very likely that further varieties of combinatorial structure will be uncovered and described in due course. At the same time, the combinatorial structures discovered and described to date are far more limited in range and scope than the phonological and morphological organization found in human languages. There remains considerable disagreement and controversy over how these findings should be interpreted (see Section 2.2, below).

The emerging picture appears to be substantial, even overwhelming variation. Even considering vocal systems alone, combinatorics of various kinds have been described across very diverse

species. In addition to the above examples of simple combinations in putty-nosed monkeys and of duality of patterning in chestnut-crowned babblers are "affixes" found in male Campbell's monkey alarm calls (Ouattara et al. 2009). These calls constitute an acoustically variable stem, which can be followed by an acoustically invariable suffix that appears to broaden the effect of the call, changing it from a specific eagle alarm to a general disturbance call, or from a specific leopard alarm to a general alert call. These are only a few examples of a large patchwork of different types of combinatorial structure, many of which have obvious and important similarities with structures observed in natural languages.

To underline the point, here is a list of terms employed in one recent taxonomy of the vocal communication systems of nonhuman animals: phonocoding, multielement calls, temporal structures, intermediate/readout structures, segmental concatenations, affixations, semantic compositionality, semantic combinatoriality, and stochastic/proportional structures (Engesser & Townsend 2019). Another taxonomy, focused on the vocal communication of nonhuman primates, uses a different array of terms: merged compounds, stochastic sequences, permutated sequences, hierarchical sequences, cross-modal sequences, limited (trivial) compositionality, and genuine (nontrivial) compositionality (Zuberbühler 2018). Comparison of these two lists reveals that (a) many of the labels used to describe these findings are adopted from language science, with suitable adjustments, and (b) there is as yet no widely agreed system of description and analysis. Furthermore, given that only a fraction of the world's species have been studied using modern methods—again, most of them only in the auditory domain—it is very likely that the full extent of the similarities with natural languages is still to be discovered. There are as yet no clear patterns linking the complexity of combinatorics to species' social and ecological conditions.

The adoption of linguistic terminology to describe the communication systems of other species is contentious. Fitch (2016, p. 98) speaks for many researchers when he writes that "the very notion that a primate communication system 'has' phonetics, phonology, morphology, syntax, semantics and pragmatics is already a *very* big assumption that carries with it a host of entailments that are in my opinion deeply problematic." The main issue is that using linguistic terminology without further comment can seem to presuppose that human natural languages and other animal communication systems are not only analogous but also homologous (either behaviorally or cognitively), and hence that natural languages are likely to be evolutionary enrichments of the communication systems of other species [e.g., Collier et al. (2014) use this approach to argue that syntax might have preceded phonology in the biological evolution of language]. In other words, linguistic terminology can seem to presuppose answers to the analogy/homology question in favor of homology. Yet the matter is in dispute.

To help make this issue graphic, one of us (T.S.P.) conducted a reductio ad absurdum experiment, replicating in bacteria the studies on putty-nosed monkeys summarized above. The model species for research in bacterial communication is *Pseudomonas aeruginosa*. Like many bacteria, *Pseudomonas aeruginosa* communicate by means of quorum sensing. Individual cells emit a specific diffusible molecule at a constant rate. They also monitor the relative density of that molecule in their immediate local environment, because the density of signaling molecules provides a proxy measure of the relative density of cells themselves. When this density reaches a specific threshold (quorum), the cells up- and downregulate specific genes, in turn generating behavior. In this way, bacterial cells can act in a collective and population density–dependent manner, such as to provide access to food. Scott-Phillips et al. (2014) and Cornforth et al. (2014) investigated whether this system is also used combinatorially. Using the same experimental design as used for key studies in putty-nosed monkeys, summarized above, the bacteria studies revealed the same pattern of results (for more recent research, see Gurney et al. 2020). Specific signals have specific effects; these effects are different from one another; and the combination of signals has effects that are different from

the sum of the effects of each signal on its own (and from control conditions of no signals). This is combinatorial communication, but it is not, we assume, homologous with natural languages—or, if it is homologous, then combinatorial communication systems derive from a common ancestor with bacteria, dating back four billion years, and their presence in humans is evolutionarily unremarkable. Furthermore, the mechanisms used to do combinatorial communication in bacteria are plainly very different from the neurally implemented processes of human speakers, not to mention whatever neural processes underpin the calls of putty-nosed monkeys. These findings are, in our view, an empirical proof of concept that when we find combinatorics in other species we cannot necessarily conclude that they are evolutionarily related to human linguistic capacities (Scott-Phillips et al. 2014). It is important to investigate whether combinatorial communication systems exist in very diverse taxa, and where we find it we should not necessarily assume evolutionary relatedness (i.e., homology). Nor should we assume either analogy or homology of the underlying mechanisms.

In the next section, we elaborate some specific dissimilarities between natural languages and the communication systems that have been discovered and described in many nonhuman animals. These dissimilarities in turn motivate the conclusion that the combinatorics of nonhuman primate vocalizations are analogous but not homologous to human languages.

2.2. Four Dissimilarities with Natural Language

The dissimilarities on which we focus are:

- Spontaneous versus volitional. Nonhuman vocal communication has more in common with sighs, grunts, laughter, and other spontaneous forms of human expression than with volitional forms such as language use.
- Finitude versus infinitude. Natural languages can in principle generate a countably infinite number of sentences, but as yet there is no evidence of infinitude in nonhuman communication.
- Literalism versus contextualism. Words and other linguistic items interact with human cognitive capacities for communication in ways that have no plausible similarities in other species.
- Narrow domains versus virtual open-endedness. Humans appear to have an open-ended range of communicative goals, in contrast to other species and apparently in violation of a basic principle of evolved communication systems.

2.2.1. Spontaneous versus volitional. In the vocal domain, humans communicate and express themselves in an extremely wide variety of ways, which are usually organized, as a first approximation, into two types (e.g., Gervais & Wilson 2005, Bryant 2020). On the one side are volitional vocalizations. These include not only verbal language use but also nonlinguistic but otherwise similar behaviors, such as impersonations of great ape vocalizations ("Ooo! Ooo!"). On the other side are spontaneous vocalizations. The most obvious examples are laughter, screams, and sighs, but these could also include the vocalizations that accompany orgasm and other moments of heightened emotion. Spontaneous vocalizations can be approximated in a volitional way, such as with fake laughter, but these volitional behaviors should not be equated with their spontaneous equivalents.

Three lines of research reveal deep similarities between the vocal communication of nonhuman primates and spontaneous vocalizations in humans. First, human spontaneous laughter is perceptually distinguishable from volitional laughter and perceptually indistinguishable from nonhuman animal vocalizations (Bryant & Aktipis 2014). That is, great ape vocalizations are auditorily similar to human spontaneous vocalizations and not human volitional vocalizations.

Second, the vocalization of emotional states employs similar brain structures across otherwise dissimilar species (Ackermann et al. 2014). Third, the acoustic similarities and differences in spontaneous laughter across primate species match established trees of species relatedness (Davila Ross et al. 2009). Collectively, these findings converge on the conclusion that spontaneous and volitional laughs are produced by different vocal and neurocognitive systems, and that there is phylogenetic continuity from nonhuman vocalizations to human spontaneous laughter (for further arguments, see, e.g., van Hooff 1972, Provine 2000, Burling 2005, Gervais & Wilson 2005, Vettin & Todt 2005, Fitch 2016). Put simply, it is very likely that the vocalizations of nonhuman primates are homologous not with language use but with (real) laughter and other spontaneous vocalizations in humans. Spontaneous expression is where homologies between human and nonhuman vocalizations are most likely to be found.

This conclusion, in turn, suggests an important research agenda that has not yet been systematically pursued. Human spontaneous vocalizations, such as laughter, could be studied for combinatorial structures using the tools and techniques that have been developed and successfully employed to study combinatorics in nonhuman vocalizations. It is possible that, through this approach, combinatorics would also be discovered in spontaneous human vocalizations. Such research would be highly relevant to the question of how to interpret findings of combinatorics in nonhumans from a linguistic perspective. More speculatively, it is possible that spontaneous vocalizations are very common in primates and as such were a trait of the last common ancestor with humans, and that these vocalizations have gradually decreased in number and type in the *Homo* lineage as other, volitional means of communication have emerged, most obviously language use.

2.2.2. Finitude versus infinitude.

Natural languages—and also many computer languages, which are human-made—can in principle generate an infinite number of sentences. This insight is often summarized with the dictum that language makes infinite use of finite means. Chomsky (2017, 2021) has called the task of explaining this infinitude the "Galilean challenge," following remarks by Galileo Galilei and other founders of modern science about the freedom and openness of human action.

Under the current incarnation of generative grammar, the Minimalist program, the apparent infinitude of natural language derives ultimately from the operation Merge, which takes as input two syntactic items and constructs from them one new syntactic item. Put simply, Merge(X,Y) forms the set $Z = \{X, Y\}$. Crucially, this new syntactic item, Z, can itself be an input into another operation of Merge, as can the output of that operation, and so on without supposed limit. This iterative process can in principle generate an unbounded array of hierarchically structured expressions. This infinitude is apparently absent in other species: All animal communication systems described so far can be analyzed as either zero-merge or one-merge systems (Rizzi 2016, Schlenker et al. 2016). Zero-merge systems are those with unitary signals, such as the famous vervet monkey alarm calls, each of which is apparently distinct from the others (Seyfarth et al. 1980). One-merge systems involve the combining, or merging, of two otherwise unitary signals. Two-merge systems take as input already-merged items (i.e., Z in the abstract formulation above). To date, no two-merge systems have been described in the communication systems used in the wild by any nonhuman, let alone higher orders of Merge that might enable the sort of recursively embedded structures that generate infinitude. Furthermore, a detailed analysis of utterances directed to the bonobo Kanzi, and his behavioral responses—Kanzi is arguably the most linguistically successful of the nonhuman apes trained on human language—reveals a significant dip in comprehension when he is asked to act on multiple objects described by a coordinate noun phrase, suggesting a failure to infer a hierarchical grammatical structure (Truswell 2017).

Thus, researchers working in the generative tradition now have a standard response to findings of combinatorics in animal communication. The response is that while these discoveries are important and fascinating on their own terms, they provide evidence only of finite combinations, not of the sort of hierarchical structures that enable infinitude, and hence not of the basic, foundational property of language that is, *ex hypothesi*, most in need of explanation (e.g., Hauser et al. 2014, Bolhuis et al. 2014, 2018, Berwick 2016, Huybregts 2020; for a reply, see, e.g., Townsend et al. 2018). Researchers working in this tradition therefore insist that "there is no evidence for human-like language syntax in any nonhuman species" (Bolhuis et al. 2014, p. 2). Nor do findings of combinations in nonhuman communication provide potential evolutionary stepping-stones toward language, because infinitude is not a gradual enrichment of finitude. (Of course, human linguistic communication entails a finite number of operations of Merge in practice: No actually produced utterance is infinitely long. The generative claim is that there is no limit to recursivity in principle.)

Why might animal signaling systems be limited to one-merge? One possibility is simply that we have not yet looked deeply enough. It is conceivable that animal signals entail higher orders of Merge, but research attention has focused on the possibility and interpretation of one-merge systems, so that is what we have so far discovered. A second possibility is that the present absence of two-merge systems is in some sense real—that for some reason there is a natural limit of one-merge on nonhuman communication. For instance, two-merge might place demands on short-term memory that one-merge does not, and hence impose a natural cognitive limit (Rizzi 2016, Zuberbühler 2020). A third possibility is that two-merge systems are unnecessary in nonhumans simply because the range of communicative domains is relatively small, making combinatorics unnecessary (Nowak et al. 2000, Scott-Phillips & Blythe 2013, Piantadosi & Fedorenko 2017). A fourth possibility is that this is simply a matter of evolutionary chance (Berwick & Chomsky 2016). Whatever the reasons, and whatever an individual researcher's views of generative grammar and the Minimalist program, it is true that there is no good evidence of infinitude in nonhuman communication systems. This is a clear dissimilarity with human communication.

2.2.3. Literalism versus contextualism.
Many research programs in the language sciences adopt a "presumption of literalness" (Wilson 2003), according to which linguistic stimuli have meanings independent of their use and interpretation. In a recent commentary on formal analyses of monkey communication, Jäger (2016, p. 111) provides a clear summary and restatement of these assumptions: The meaning of a word (or sentence, or utterance) is distinct from its interpretation, sentence meanings derive from the meanings of their components and the ways they are combined, and the interpretation of a sentence when it is used in discourse is inferred from the application of pragmatic principles to the meaning of the sentence. In other words, it is possible to fully encode meaning linguistically, and pragmatic enrichment is optional rather than obligatory. This is a convenient and often productive idealization, but it is not an accurate characterization of the cognitive process by which linguistic meaning is determined in ordinary use. In this subsection, we summarize what cognitive pragmatics, psycholinguistics, and the neuroscience of language have uncovered about the process of linguistic comprehension, and how it differs from the idealized presumption of literalness. There is no evidence that any nonhuman animal interprets communicative stimuli in anything like the same way.

Consider an utterance such as "I have a temperature," said by Mary to Peter in response to Peter's suggestion that Mary should visit her aunt in the hospital (for a detailed analysis of this and cognate examples, see, e.g., Wilson 2003). Interpreted in supposedly literal terms, *I have a temperature* is trivially true and hence not worth saying. Interpreted in context, it is a suggestion that Mary should probably not visit the hospital at the present time. How is this interpretation

derived? According to approaches based on the presumption of literalness, the process must be something like the following serial procedure: The literal meaning is derived, assessed as trivially true and hence irrelevant, and then enriched with pragmatic assumptions. If this is indeed how linguistic comprehension works, then one prediction that follows immediately is that metaphorical utterances, along with irony, indirect speech, implicature, and indeed all supposedly nonliteral utterances, should take longer to process, and be more prone to error, than supposedly more-literal cases. Yet there are simply no good data to support this prediction. On the contrary, a great deal of psycholinguistic data speak against it (Gibbs & Colston 2012).

Observational and neuroscientific data are far more consistent with a contextualist picture, in which inference of what is explicit and inference of what is implicit are computed not serially but in parallel, with each influencing the other in a dynamic process of "mutual adjustment" (Wilson 2003, Wilson & Carston 2007; for neuroscientific evidence, see, e.g., Spotorno et al. 2012, Vanlangendonck et al. 2018, Paunov et al. 2019). In this picture, the linguistic decoding of *I have a temperature* activates Peter's knowledge about temperatures, hospital visits, and the possible connections between them, causing Peter to adjust, in parallel, his representations of (*a*) how the words have been used (i.e., the different ways in which Mary could be using the word *temperature*) and (*b*) what the speaker means (i.e., the different ways in which Mary might be responding to Peter's suggestion that she visit the hospital). Representations *a* and *b* are each adjusted to fit the other, until an interpretation is determined that Peter finds relevant enough to terminate the comprehension process. In consequence, Peter arrives at an interpretation of Mary's words in which the spoken word *temperature* does not map onto its supposed literal meaning at all. Instead, Peter interprets *temperature* as mapping onto a concept that is temporary and ad hoc, namely a temperature high enough to make it inadvisable for Mary to visit her aunt in the hospital—even though this concept is not what the word *temperature* typically encodes. Simultaneously, Peter arrives at a complementary interpretation of what is meant, namely that Mary does not feel able to visit her aunt. Humans have specialized neurocognitive processes for precisely this task of parallel mutual adjustment (Hagoort 2019). Moreover, all linguistically encoded concepts are subject to this process of enrichment and adjustment, albeit to a greater degree in some cases than others. The limiting case of relatively little enrichment should not be confused with the idealization of wholly unenriched, literal meanings (Wilson & Carston 2007).

As stated above, there is no evidence in any nonhuman species that the interpretation of what is meant entails a process of parallel mutual adjustment in which how the stimuli are used and what they mean are adjusted in parallel, with the process terminating once the audience converges on a pair of interpretations that satisfy prior expectations of relevance. Again, for many research purposes the process of linguistic comprehension (not to mention production) can be conveniently idealized in terms of literal meanings and independent pragmatic enrichments, but in making comparisons across species one should not confuse this idealization with the underlying cognitive capacities. The existence of specialized processes for linguistic interpretation—and likely linguistic production, too—is an important dissimilarity between human and nonhuman cognition.

2.2.4. Narrow domains versus virtual open-endedness. Living things communicate in a great variety of ways, but human communication is open-ended in a way that is clearly distinctive. Not only is there language use, but also there are points, nods, winks, and other behaviors that, although not linguistic, are still conventionalized, along with many ad hoc, improvised behaviors, such as a small hand gesture used to visually park a topic of ongoing conversation. In contrast, all nonhuman communication systems appear to be limited to finite and specific domains. Bees, for instance, use combinatorics to communicate about the location of flowers and the quantity of their nectar, but apparently little or nothing else. There is still much still to discover about animal

communication; from a comparative perspective it is important to note that the gestural communication of nonhuman great apes is apparently more diverse and flexible than most other modes of animal communication, but even here the range and scope are plainly limited relative to the human case.

There are good evolutionary reasons for communication systems to be tied to narrow domains. For communication to be stable, it must be beneficial, on average, to both communicator and audience (e.g., Maynard Smith & Harper 2003, Searcy & Nowicki 2005, Scott-Phillips et al. 2012). This does not imply that communication is always of mutual benefit or that deception never occurs, but it does imply that communication must be sufficiently beneficial, sufficiently often, for both parties, because otherwise it would collapse. In some cases, the mutual benefit derives from genetic relatedness, such as with ant pheromones or bee dance. In other cases, the mutual benefit derives from direct fitness effects on communicator and audience. Whatever the reasons, evolutionary dynamics leash communication to relatively narrow domains of statistical mutual benefit.

Human communication—and apparently only human communication—appears to be in flagrant violation of this limitation. Its range is certainly not restricted to any particular topic: Humans can communicate about potentially anything. Moreover, humans frequently communicate about phenomena for which no directly observable evidence could ever be provided, as in statements about past or future events. The gestural communication of nonhuman great apes is apparently more diverse and flexible than most other modes of animal communication (see Section 3.1, below), but its scope is still clearly limited relative to that of humans.

We have argued elsewhere that this apparent violation of a basic principle of evolutionary theory is fundamental to explaining the apparent open-endedness of human communication, and we have developed an answer based on the distinctively metapsychological structure of human communication, which allows the domain of communication to be narrow and specific but virtually open-ended (Heintz & Scott-Phillips 2023). Other researchers aiming to explain the open-endedness of human communication emphasize specifically combinatorial aspects of natural language (see Section 2.2.2, above). Whatever the merits of these or any other specific explanation, the difference in communicative systems is plain. Humans have an open-ended range of goals in communication and a large and wide range of means with which to satisfy those goals; and while some forms of nonhuman communication are more open-ended and flexible than others, none appears to exhibit this full open-ended range and diversity.

2.3. Similarities and Dissimilarities: Conclusion

Collectively, these dissimilarities strongly suggest that while the vocal communication systems observed in nonhuman species are analogous to natural languages and human language use in interesting and important ways, they are not homologous. Nor do they support the hypothesis that some underlying cognitive mechanisms have independently evolved in all communicating organisms. Depending on the trait in question, data relevant to the homology/analogy question can come in many forms: phylogenetic, genetic, anatomic, neuroscientific, behavioral, and so on. The key measures are similarity and difference with the focal trait. The data reviewed above reveal that, while communication systems in nonhumans share some surface similarities with natural language, there are also many important dissimilarities, which collectively constitute strong evidence against homology. Combinatorial communication systems in nonhumans have interesting surface similarities with natural language, but stronger conclusions are unwarranted.

In searching for relevant homologies, a more promising approach is to focus on social cognition. In the next section, we describe how there may be important evolutionary homologies to

be identified and described particularly in the domain of attention manipulation, with relatively small differences between humans and other great apes.

3. INTENTIONS AND MEANING

3.1. A Puzzle of Great Ape Gesture

Arguably the most intriguing thing about great ape interactions is that they often inform one another in ways that can seem very human. Here are two specific examples. First, male bonobos sometimes use a beckoning behavior to garner attention and encourage others to join them. An arm is stretched out toward the audience, followed by a sideways, sweeping movement of the arm toward the self, in a way that is strikingly similar in form to human beckoning. This behavior has been reported in the wild and, more recently, documented on video as part of systematic observational studies of bonobos living in near-natural environmental conditions (Genty & Zuberbühler 2004). Beckoning is used especially (but not only) to initiate sexual intercourse, and as such it is often preceded by exaggerated presentation of an erect penis. A second well-documented example is gestures used to request grooming, such as a big, loud scratch across one's own chest (Hobaiter & Byrne 2014). In both examples—and to be clear, these examples are but two of many—the behavior is repeated if no suitable reaction follows. Great apes use a wide array of gestures to interact with one another, many of which look very much like the sort of spontaneous nonverbal behaviors that humans use to communicate with one another (for recent reviews, see Byrne et al. 2017, Tomasello & Call 2019). These gestures are used flexibly, they take account of the audience's attentional state, and individuals will readily substitute one gesture for another when confronting difficulty in comprehension. For these reasons and others, great ape gesture is commonly called intentional. How should we interpret such behaviors from a linguistic perspective?

The relevant theoretical literature on intentions and meaning has its most important origins in the work of Grice (1957, 1989), who characterized speaker meaning as, at bottom, a psychological phenomenon predicated on expressing and recognizing intentions. His most crucial insight was that meaning is "auto-deictic": that when speakers "mean" something, they do not simply intend to achieve a particular effect in the audience. It is also part of meaning that speakers achieve this effect by revealing, or pointing to, their own intention to achieve that effect. In other words, not only is meaningful behavior intentional, it is overtly intentional: The intention itself is made overt. Eating, for instance, is an intentional behavior; but sometimes we eat in an exaggerated way, to express to others that the food is tasty, revolting, generous, or fancy. When we do so, we intend that the audience learn about the food, and we satisfy this intention by showing that this is our intention ('Look! I am trying to show that the food is tasty!'). Communication by means of overt intentionality is often called Gricean communication or, in some theoretical frameworks, ostensive communication.

This auto-deictic quality is absent from many analyses of nonhuman primate communication (e.g., Moore 2017). The following passage summarizes the approach of one productive research agenda:

> "Meaning" is a loaded term when discussing animal communication. Normally, in order to avoid unwarranted imputation of goals to signallers, biologists describe signals by their function: the effects they produce on audiences and the fitness benefits of these effects for the signaller.... But since the intentionality of ape gesture has been robustly established, it is appropriate to ask what signalers *mean*: what effects do they *want* to produce? (Byrne et al. 2017, p. 764)

This use of "meaning" is weaker than the Gricean notion summarized above. Specifically, it removes the auto-deictic aspect that Grice paid special attention to (Scott-Phillips 2015a, pp. 22–23;

Warren & Call 2022). This focus is instead only on the apparent goal. The research agenda that follows from this focus is to look for behavioral evidence of "apparently satisfactory outcomes" (Byrne et al. 2017, p. 764), by observing how audiences react to gestures and, in particular, what reactions cause the gesturer to cease performing the gesture (e.g., Cartmill & Byrne 2010, Hobaiter & Byrne 2014).

In contrast, many human scientists naturally focus on the full expressive richness of Gricean communication, which extends well beyond language use. Pointing is arguably the most obvious and salient example of nonlinguistic Gricean communication, but in fact anything humans can do they can do in a Gricean or ostensive way, as in the above example of eating food in an overt way to express an opinion about it. The complete range of Gricean communication is a matter for empirical discovery and plausibly includes some behaviors that are not always recognized as communicative at all, such as punishment, artistic expression, the many diverse forms of teaching that humans use, subtle movements that connect dance partners, and others (Heintz & Scott-Phillips 2023). As stated above in Section 2.2.4, this massive and open expressive range is not observed in other species.

So, which way do we turn? Is great ape gesture evidence of the same, Gricean capacity of mind that provides the foundations for a great many human behaviors, including language use? Or is it something else, and if so, then what? However we answer these questions, there are difficult follow-ups (for relevant debate, see, e.g., Moore 2016; Scott-Phillips 2015a, 2016). If we say that yes, great apes are Gricean communicators, then why is great ape communication still not nearly as rich and open-ended as human expression? Turning the other way, if we say no, great apes are not Gricean communicators, then on what grounds exactly is this judgment made? Many great ape behaviors look like the sort of thing that humans sometimes do in communication.

The puzzle, then, is how to describe great ape gesture in a way that accounts for both (*a*) its behavioral similarity with some forms of human communication and (*b*) its apparently limited range and scope relative to the human case. The next section describes one way to proceed.

3.2. Layers of Attention Manipulation

Elsewhere we have described the evolution of human communication in terms of the evolution of new ways of manipulating others' attention, and of reading others' attempts to manipulate attention (Scott-Phillips 2015b, Heintz et al. 2016, Heintz & Scott-Phillips 2023). Here we enrich this perspective further, to directly address the puzzle of great ape gesture described above.

Figure 1 presents five embedded subsets, each a particular means of manipulating attention. This "special case of" approach is in clear contrast to existing approaches based on a supposed hierarchy of levels of intentionality. It also meets a growing demand for analyses that go beneath the surface of behavior to consider underlying cognitive processes (e.g., Graham et al. 2020).

To describe these subsets we begin in the middle and work inward. The Gricean subset we have already characterized above: the intentional manipulation of attention toward one's own informative intentions (see Section 3.1). Other researchers—particularly Wilson & Sperber (2012)—have since developed this idea in more cognitively precise ways than Grice himself did, but it was Grice who first emphasized the auto-deictic nature of human communication. Lewisian is Gricean communication by means of convention (following Lewis 1969). Conventions are commonly known solutions to recurrent coordination problems, and in the case of communication the recurrent coordination problem is identification of the speaker's informative intention. A great many instances of Gricean communication, such as nodding, winking, and pointing, entail communicative conventions so defined. At the same time, not all Gricean communication is conventional. Humans can improvise novel behaviors when necessary, which is sometimes called pantomime (e.g., Zlatev

Figure 1

Subsets of means of manipulating others' attention. (①) Washburnian: intentional manipulation of attention. (②) Ladyginian: intentional manipulation of attention toward one's own intentions. (③) Gricean: intentional manipulation of attention toward one's own informative intentions. (④) Lewisian: intentional manipulation of attention toward one's own informative intentions, by means of convention. (⑤) Saussurian: intentional manipulation of attention toward one's own informative intentions, by means of convention, where conventions are organized in structured sets. These structured sets are called languages. Note that the distinctions between these subsets are graded (more or less) rather than categorical (this or that). This taxonomy in turn allows us to advance the hypothesis that great ape gestural communication is best characterized as falling within subset ②, Ladyginian.

et al. 2020). Imagine attempting to communicate with someone in a cultural context deeply unfamiliar to you, such that you have no shared language or knowledge of other communicative conventions. Under these circumstances, pantomime may be necessary. Such communication is Gricean, but if it involves no conventions, it is not Lewisian. A further subset of Lewisian communication is when the conventions in question are (self-)organized in structured networks that are stable but also dynamic and constantly subject to change. These sets are commonly called languages, and the conventions that constitute them are sometimes called constructions. We call this layer Saussurian, following the emphasis that Ferdinand de Saussure placed on structure in linguistic description and analysis.

At the other extreme, the outermost subset includes all instances of the intentional manipulation of attention. We assume that this subset extends beyond humans and could plausibly include, for instance, gray squirrels, which have been shown to modulate their caching behavior as a function of the presence of onlookers (e.g., moving a cache when the onlooker leaves) (Leaver et al. 2007), and ravens, which have been shown to guard their caches against discovery, taking into account other ravens' possible knowledge of the cache (Bugnyar et al. 2016). We call this subset Washburnian after Margaret Floy Washburn (1908), who pioneered the study of social cognition in a wide variety of nonhuman animals, at a time when more-behaviorist approaches were coming into vogue. (Washburn was also the first woman to obtain a PhD in psychology, at Cornell in 1894.)

The key novelty in our analysis is specification of the second innermost subset (for a somewhat similar analysis, see Warren & Call 2022). Here, individuals intentionally manipulate others' attention toward evidence of the intended outcome, which could be, for instance, grooming, play, or sex. This is possible if observers can identify such intentions on the basis of the observed behavior. We name this subset Ladyginian after Nadezhda Ladygina-Kohts (born Nadezhda Ladygina).

Ladyginian: refers to intentional manipulation of attention toward one's own intentions; plausibly a good characterization of great ape interaction

Ladygina-Kohts was a pioneer in the comparative study of great ape social cognition. Foreshadowing some highly productive research agendas pursued in subsequent decades, she was one of the first scientists to document in a systematic way the similarities and differences between the expressive behaviors of humans and chimpanzees (Ladygina-Kohts & de Waal 2002; for further historical context, see Cartmill & Hobaiter 2019).

The difference between Ladyginian and Washburnian is that whereas Washburnian behavior is intentional toward any end, Ladyginian behavior intentionally reveals intentions (to be groomed, to play, etc.). The difference between Ladyginian and Gricean is that Gricean behavior intentionally reveals not simply intentions but specifically informative intentions. Or, to put the essential difference between Ladyginian and Gricean very concisely, one is auto-deictic and the other is not (see Section 3.1). We suggest that most great ape gesture may be best characterized as Ladyginian but not Gricean.

This suggestion in turn generates plausible new explanations of some important empirical findings. First, human infants will object if they are misunderstood—in other words, if their informative intention is not satisfied—even when their material goals in communication were nevertheless satisfied (Grosse et al. 2010). There is no evidence of any similar dispositions in great apes. This difference between species is exactly the pattern we should expect if changing others' mental states is, in one mode of interaction, a means to an end (Ladyginian) and, in the other, the end itself (Gricean). As a second, more speculative example, great ape gesture tends to have a dyadic rather than triadic character. That is, great ape gestures attract the audience's attention to the self, for immediate social goals such as play or sex, rather than to some other, third entity, such as a distant object (for a review, see, e.g., Pika et al. 2005). This pattern is not absolute—great ape interaction does seem to be triadic under some circumstances (e.g., Lyn et al. 2014, Bohn et al. 2015)—but human communication is plainly more triadic, more often, than great ape interactions. Again, this pattern might be explained by the distinction between Ladyginian and Gricean. In Ladyginian interaction attention is drawn directly to the focal individual's behavioral intentions, whereas in Gricean interaction attention is drawn to the focal individual's informative intentions, which are only in turn about anything at all. In this way, Gricean communication opens up or unleashes expression on a massive scale (Heintz & Scott-Phillips 2023).

Two final points about great ape gesture: First, our suggestions here about the empirical nature of great ape gesture are potentially convergent with some other analyses (see, in particular, Moore 2016, 2017). However, those previous analyses maintain the label Gricean for great ape gesture. We think that label conflates two distinct modes of interaction—one auto-deictic and the other not—that should be kept separate, as in **Figure 1**. Second, behaviors that are Ladyginian but not Gricean can be interpreted by the intended audiences with ordinary cognitive capacities of mind reading, but this is not so for the interpretation of specifically Gricean communication, which depends on more narrowly specialized forms of social cognition, in particular audience presumptions of relevance that are triggered by ostensive behavior (Sperber & Wilson 2002; Heintz & Scott-Phillips 2023, section 4). In humans, these specialized capacities are part of the ordinarily developing cognitive phenotype (Heintz & Scott-Phillips 2023).

4. CONCLUSION: MEANING AND GRAMMAR HAVE DIFFERENT ETIOLOGIES

Reviewing the literature on animal communication from a linguistic perspective, we have argued for a complementary pair of conclusions. In Section 2, we argue that many animal communication systems, in particular those in the vocal domain, are analogous but not homologous to natural languages. Does this mean that the description of animal communication systems is largely

irrelevant to understanding natural languages? Not necessarily. As they emerge, change, and evolve, languages are shaped by the various biases, needs, and capacities of language learners and language users. These various dispositions can be of many different kinds. Some are clearly general aspects of cognition, such as the distinction between objects and actions, which helps us navigate the physical world and hence causes languages to tend to differentiate nouns and verbs. Other dispositions might be specifically syntactic in character (for three different approaches to this issue, see, e.g., Sperber & Origgi 2010, Berwick & Chomsky 2016, Kirby 2017). What animal communication research reveals is the extent to which some dispositions, especially capacities to combine perceptually distinct items, are shared with other species (Zuidema & de Boer 2018).

In Section 3, we argue that there is evolutionary continuity between humans and other great apes in another domain, namely the means by which individuals can manipulate others' attention. In fact, the cognitive differences between humans and other great apes in this domain may be small and subtle, albeit with large consequences. We identified a key distinction between making manifest intentions in general (Ladyginian) and making manifest specifically informative intentions (Gricean). Elsewhere we have argued that this shift derives from the distinctive nature of human social ecologies, which generate selection for specialized cognitive capacities of production and interpretation (Heintz & Scott-Phillips 2023). Linguistic expression exploits these Gricean foundations (e.g., Levinson 2006, Scott-Phillips 2015b, Moore 2017, Seyfarth & Cheney 2018).

This pair of conclusions, in turn, suggests that meaning and grammar have different etiologies (see also Tomasello 2003). Meaning derives from a species-distinct capacity to express and recognize informative intentions. This capacity is often called Gricean, and in our view it is described in its most cognitively plausible form in Relevance Theory (Wilson & Sperber 2012, Clark 2013). The origins of human language can be found in the evolution of these capacities (Heintz & Scott-Phillips 2023). Grammars, meanwhile, are descriptions of sets of communication conventions: in effect, tools employed in the service of Gricean communication (**Figure 1**). These tools are shaped by many factors as they evolve and change over time. Some of these factors are cognitive dispositions and capacities that are clearly shared with other species.

SUMMARY POINTS

1. Combinatorial communication is far more widespread in the natural world than was known even 20 years ago.
2. These combinatorial systems are analogous but not homologous to human linguistic capacities.
3. There are evolutionary homologies between humans and other great apes in cognitive capacities of attention manipulation.
4. Human communication, including language use, is a specialized form of attention manipulation.

FUTURE ISSUES

1. What is the full extent of combinatorial communication in the natural world?
2. Is there combinatoriality in human spontaneous vocalizations, such as laughter?
3. How can Gricean capacities evolve from Ladyginian capacities? Under what ecological conditions is this most likely to occur?

4. What specific social cognitive capacities explain key similarities and differences in social interaction between humans and other great apes?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

LITERATURE CITED

Ackermann H, Hage SR, Ziegler W. 2014. Brain mechanisms of acoustic communication in humans and nonhuman primates: an evolutionary perspective. *Behav. Brain Sci.* 37(6):529–46

Anderson SR. 2006. *Doctor Dolittle's Delusion: Animals and the Uniqueness of Human Language*. New Haven, CT: Yale Univ. Press

Arnold K, Zuberbühler K. 2006. Semantic combinations in primate calls. *Nature* 441(7091):303

Arnold K, Zuberbühler K. 2008. Meaningful call combinations in a non-human primate. *Curr. Biol.* 18(5):R202–3

Berwick RC. 2016. Monkey business. *Theor. Linguist.* 42(1/2):91–95

Berwick RC, Chomsky N. 2016. *Why Only Us? Language and Evolution*. Cambridge, MA: MIT Press

Berwick RC, Okanoya K, Beckers GJL, Bolhuis JJ. 2011. Songs to syntax: the linguistics of birdsong. *Trends Cogn. Sci.* 15(3):113–21

Bohn M, Call J, Tomasello M. 2015. Communication about absent entities in great apes and human infants. *Cognition* 145:63–72

Bohn M, Frank MC. 2019. The pervasive role of pragmatics in early language. *Annu. Rev. Dev. Psychol.* 1:223–49

Bolhuis JJ, Beckers GJ, Huybregts MA, Berwick RC, Everaert MB. 2018. Meaningful syntactic structure in songbird vocalizations? *PLOS Biol.* 16(6):e2005157

Bolhuis JJ, Okanoya K, Scharff C. 2010. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat. Rev. Neurosci.* 11(11):747–59

Bolhuis JJ, Tattersall I, Chomsky N, Berwick RC. 2014. How could language have evolved? *PLOS Biol.* 12(8):e1001934

Bryant GA. 2020. Evolution, structure, and functions of human laughter. In *Handbook of Communication Science and Biology*, ed. K Floyd, R Weber, pp. 63–77. London: Routledge

Bryant GA, Aktipis CA. 2014. The animal nature of spontaneous human laughter. *Evol. Hum. Behav.* 35(4):327–35

Bugnyar T, Reber SA, Buckner C. 2016. Ravens attribute visual access to unseen competitors. *Nat. Commun.* 7:10506

Burling R. 2005. *The Talking Ape: How Language Evolved*. Oxford, UK: Oxford Univ. Press

Byrne RW, Cartmill E, Genty E, Graham KE, Hobaiter C, Tanner J. 2017. Great ape gestures: intentional communication with a rich set of innate signals. *Anim. Cogn.* 20(4):755–69

Cartmill EA, Byrne RW. 2010. Semantics of primate gestures: intentional meanings of orangutan gestures. *Anim. Cogn.* 13(6):793–804

Presents a skeptical analysis on combinatorial communication in songbirds, from the perspective of generative grammar.

Provides an overview of the St Andrews approach to meaning in great ape gesture.

Cartmill EA, Hobaiter C. 2019. Developmental perspectives on primate gesture: 100 years in the making. *Anim. Cogn.* 22(4):453–59

Chomsky N. 2017. The Galilean challenge. *Inference* 3(1). **https://doi.org/10.37282/991819.17.1**

Chomsky N. 2021. Linguistics then and now: some personal reflections. *Annu. Rev. Linguist.* 7:1–11

Clark B. 2013. *Relevance Theory*. Cambridge, UK: Cambridge Univ. Press

Collier K, Bickel B, van Schaik CP, Manser MB, Townsend SW. 2014. Language evolution: syntax before phonology? *Proc. R. Soc. B* 281(1788):20140263

Cornforth D, Popat R, McNally L, Gurney J, Scott-Phillips T, et al. 2014. Combinatorial quorum-sensing communication allows bacteria to resolve physical and social uncertainty. *PNAS* 111(11):4280–84

Davila Ross M, Owren MJ, Zimmermann E. 2009. Reconstructing the evolution of laughter in great apes and humans. *Curr. Biol.* 19(13):1106–11

de Boer B, Sandler W, Kirby S. 2012. New perspectives on duality of patterning: introduction to the special issue. *Lang. Cogn.* 4(4):251–59

Dediu D, de Boer B. 2016. Language evolution needs its own journal. *J. Lang. Evol.* 1(1):1–6

Engesser S, Crane JM, Savage JL, Russell AF, Townsend SW. 2015. Experimental evidence for phonemic contrasts in a nonhuman vocal system. *PLOS Biol.* 13(6):e1002171

Engesser S, Fitch WT. 2021. Babbler phonology and combinatorial systems. *Inference* 6(2). **https://doi.org/10.37282/991819.21.33**

Engesser S, Holub JL, O'Neill LG, Russell AF, Townsend SW. 2019. Chestnut-crowned babbler calls are composed of meaningless shared building blocks. *PNAS* 116(39):19579–84

Engesser S, Townsend SW. 2019. Combinatoriality in the vocal systems of nonhuman animals. *Wiley Interdiscip. Rev. Cogn. Sci.* 10(4):e1493

Fischer J. 2017. Playback experiments. In *The International Encyclopedia of Primatology*, ed. A Fuentes. New York: Wiley. **https://doi.org/10.1002/9781119179313.wbprim0140**

Fitch WT. 2016. Why formal semantics and primate communication make strange bedfellows. *Theor. Linguist.* 42(1/2):97–109

Fitch WT. 2018. The biology and evolution of speech: a comparative analysis. *Annu. Rev. Linguist.* 4:255–79

Fitch WT, Friederici AD. 2012. Artificial grammar learning meets formal language theory: an overview. *Philos. Trans. R. Soc. B* 367(1598):1933–55

Genty E, Zuberbühler K. 2014. Spatial reference in a bonobo gesture. *Curr. Biol.* 24(14):1601–5

Gervais M, Wilson DS. 2005. The evolution and functions of laughter and humor: a synthetic approach. *Q. Rev. Biol.* 80(4):395–430

Ghazanfar AA, Rendall D. 2008. Evolution of human vocal production. *Curr. Biol.* 18(11):R457–60

Gibbs RW Jr., Colston HL. 2012. *Interpreting Figurative Meaning*. Cambridge, UK: Cambridge Univ. Press

Graham KE, Wilke C, Lahiff NJ, Slocombe KE. 2020. Scratching beneath the surface: intentionality in great ape signal production. *Philos. Trans. R. Soc. B* 375(1789):20180403

Grice HP. 1957. Meaning. *Philos. Rev.* 66(3):377–88

Grice HP. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard Univ. Press

Grosse G, Behne T, Carpenter M, Tomasello M. 2010. Infants communicate in order to be understood. *Dev. Psychol.* 46(6):1710–22

Gurney J, Azimi S, Brown SP, Diggle SP. 2020. Combinatorial quorum sensing in *Pseudomonas aeruginosa* allows for novel cheating strategies. *Microbiology* 166(8):777–84

Hagoort P. 2019. The neurobiology of language beyond single-word processing. *Science* 366(6461):55–58

Hauser MD, Yang C, Berwick RC, Tattersall I, Ryan MJ, et al. 2014. The mystery of language evolution. *Front. Psychol.* 5:401

Heesen R, Hobaiter C, Ferrer-i-Cancho R, Semple S. 2019. Linguistic laws in chimpanzee gestural communication. *Proc. R. Soc. B* 286(1896):20182900

Heintz C, Karabegovic M, Molnar A. 2016. The co-evolution of honesty and strategic vigilance. *Front. Psychol.* 7:1503

Heintz C, Scott-Phillips T. 2023. Expression unleashed: the evolutionary and cognitive foundations of human communication. *Behav. Brain Sci.* In press

Hobaiter C, Byrne RW. 2014. The meanings of chimpanzee gestures. *Curr. Biol.* 24(14):1596–1600

Describes experiments showing songbird vocalizations with the features commonly known as duality of patterning.

Describes how and why humans, and only humans, evolved capacities necessary for open-ended communication.

Huybregts MA. 2020. Babbling birds. *Inference* 5(3). **https://doi.org/10.37282/991819.20.39**

Hurford JR. 2007. *The Origins of Meaning: Language in the Light of Evolution*, Vol. 1. Oxford, UK: Oxford Univ. Press

Hurford JR. 2012. *The Origins of Grammar: Language in the Light of Evolution*, Vol. 2. Oxford, UK: Oxford Univ. Press

Kirby S. 2017. Culture and biology in the origins of linguistic structure. *Psychon. Bull. Rev.* 24(1):118–37

Jäger G. 2016. Grice, Occam, Darwin. *Theor. Linguist.* 42(1/2):111–15

Ladygina-Kohts NN, de Waal FBM. 2002. *Infant Chimpanzee and Human Child: A Classic 1935 Comparative Study of Ape Emotions and Intelligence*. Oxford, UK: Oxford Univ. Press

Leaver LA, Hopewell L, Caldwell C, Mallarky L. 2007. Audience effects on food caching in grey squirrels (*Sciurus carolinensis*): evidence for pilferage avoidance strategies. *Anim. Cogn.* 10(1):23–27

Leroux M, Townsend SW. 2020. Call combinations in great apes and the evolution of syntax. *Anim. Behav. Cogn.* 7(2):131–39

Levinson SC. 2006. On the human 'interactional engine'. In *Roots of Human Sociality: Culture, Cognition and Interaction*, ed. NJ Enfield, SC Levinson, pp. 39–69. Oxford, UK: Berg

Lewis D. 1969. *Convention*. Cambridge, MA: Harvard Univ. Press

Lyn H, Russell JL, Leavens DA, Bard KA, Boysen ST, et al. 2014. Apes communicate about absent and displaced objects: Methodology matters. *Anim. Cogn.* 17(1):85–94

Maynard Smith J, Harper D. 2003. *Animal Signals*. Oxford, UK: Oxford Univ. Press

Moore R. 2016. Meaning and ostension in great ape gestural communication. *Anim. Cogn.* 19(1):223–31

Moore R. 2017. Convergent minds: ostension, inference and Grice's third clause. *Interface Focus* 7:20160107

Munz T. 2005. The bee battles: Karl von Frisch, Adrian Wenner and the honey bee dance language controversy. *J. Hist. Biol.* 38(3):535–70

Nowak MA, Plotkin JB, Jansen VAA. 2000. The evolution of syntactic communication. *Nature* 404(6777):495–98

Ouattara K, Lemasson A, Zuberbühler K. 2009. Campbell's monkeys use affixation to alter call meaning. *PLOS ONE* 4(11):e7808

Paunov AM, Blank IA, Fedorenko E. 2019. Functionally distinct language and Theory of Mind networks are synchronized at rest and during language comprehension. *J. Neurophysiol.* 121(4):1244–65

Pepperberg IM. 2017. Animal language studies: What happened? *Psychon. Bull. Rev.* 24(1):181–85

Piantadosi ST, Fedorenko E. 2017. Infinitely productive language can arise from chance under communicative pressure. *J. Lang. Evol.* 2(2):141–47

Pika S, Liebal K, Call J, Tomasello M. 2005. Gestural communication of apes. *Gesture* 5(1/2):41–56

Provine RR. 2000. *Laughter: A Scientific Investigation*. New York: Penguin

Radick G. 2007. *The Simian Tongue: The Long Debate About Animal Language*. Berkeley: Univ. Calif. Press

Rendall D. 2021. Aping language: historical perspectives on the quest for semantics, syntax, and other rarefied properties of human language in the communication of primates and other animals. *Front. Psychol.* 12:675172

Rizzi L. 2016. Monkey morpho-syntax and Merge-based systems. *Theor. Linguist.* 42(1/2):139–45

Sandler W. 2012. The phonological organization of sign languages. *Lang. Linguist. Compass* 6(3):162–82

Scarantino A, Clay Z. 2015. Contextually variable signals can be functionally referential. *Anim. Behav.* 100:e1–8

Schlenker P, Chemla E, Schel AM, Fuller J, Gautier JP, et al. 2016. Formal monkey linguistics. *Theor. Linguist.* 42(1/2):1–90

Scott-Phillips T. 2015a. Meaning in animal and human communication. *Anim. Cogn.* 18(3):801–5

Scott-Phillips T. 2015b. *Speaking Our Minds: Why Human Communication Is Different and How Language Evolved to Make It Special*. London: Palgrave Macmillan

Scott-Phillips T. 2016. Meaning in great ape communication: summarising the debate. *Anim. Cogn.* 19(1):233–38

Scott-Phillips T. 2017. Pragmatics and the aims of language evolution. *Psychon. Bull. Rev.* 24(1):186–89

Scott-Phillips T, Blythe RA. 2013. Why is combinatorial communication rare in the natural world, and why is language an exception to this trend? *J. R. Soc. Interface* 10(88):20130520

Presents a fascinating history of playback experiments.

Provides an overview of a research program applying tools of formal semantics to nonhuman primate communication.

Scott-Phillips T, Blythe R, Gardner A, West S. 2012. How do communication systems emerge? *Proc. R. Soc. B* 279:1943–49

Scott-Phillips T, Diggle S, Gurney J, Ivens A, Popat R. 2014. Combinatorial communication in bacteria: implications for the origins of linguistic generativity. *PLOS ONE* **9(4):e95929**

Scott-Phillips T, Kirby S. 2010. Language evolution in the laboratory. *Trends Cogn. Sci.* 14(9):411–17

Scott-Phillips T, Kirby S. 2013. Information, influence and inference in language evolution. In *Animal Communication Theory: Information and Influence*, ed. U Stegmann, pp. 421–42. Cambridge, UK: Cambridge Univ. Press

Searcy WA, Nowicki S. 2005. *The Evolution of Animal Communication*. Princeton, NJ: Princeton Univ. Press

Seyfarth RM, Cheney DL. 2003. Signalers and receivers in animal communication. *Annu. Rev. Psychol.* 54:145–73

Seyfarth RM, Cheney DL. 2018. *The Social Origins of Language*. Princeton, NJ: Princeton Univ. Press

Seyfarth RM, Cheney DL, Marler P. 1980. Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim. Behav.* 28:1070–94

Sperber D, Origgi G. 2010. A pragmatic account of the origin of language. In *The Evolution of Human Language: Biolinguistic Perspectives*, ed. RK Larson, V Déprez, H Yamakido, pp. 124–31. Cambridge, UK: Cambridge Univ. Press

Sperber D, Wilson D. 2002. Pragmatics, modularity and mind-reading. *Mind Lang.* 17(1/2):3–23

Spotorno N, Koun E, Prado J, van der Henst JB, Noveck IA. 2012. Neural evidence that utterance-processing entails mentalizing: the case of irony. *NeuroImage* 63(1):25–39

Suzuki TN, Zuberbühler K. 2019. Animal syntax. *Curr. Biol.* 29(14):R669–71

ten Cate C. 2017. Assessing the uniqueness of language: Animal grammatical abilities take center stage. *Psychon. Bull. Rev.* 24(1):91–96

Tomasello M. 2003. On the different origins of symbols and grammar. In *Language Evolution*, ed. S Kirby, M Christiansen, pp. 94–110. Oxford, UK: Oxford Univ. Press

Tomasello M, Call J. 2019. Thirty years of great ape gestures. *Anim. Cogn.* 22(4):461–69

Tomasello M, Carpenter M, Liszkowski U. 2007. A new look at infant pointing. *Child Dev.* 78(3):705–22

Townsend SW, Engesser S, Stoll S, Zuberbühler K, Bickel B. 2018. Compositionality in animals and humans. *PLOS Biol.* 16(8):e2006425

Truswell R. 2017. Dendrophobia in bonobo comprehension of spoken English. *Mind Lang.* 32(4):395–415

van Hooff JA. 1972. A comparative approach to the phylogeny of laughter and smiling. In *Nonverbal Communication*, ed. RA Hinde, pp. 209–41. Cambridge, UK: Cambridge Univ. Press

Vanlangendonck F, Willems RM, Hagoort P. 2018. Taking common ground into account: specifying the role of the mentalizing network in communicative language production. *PLOS ONE* 13(10):e0202943

Vettin J, Todt D. 2005. Human laughter, social play, and play vocalizations of nonhuman primates: an evolutionary approach. *Behaviour* 142:217–40

Warren E, Call J. 2022. Inferential communication: bridging the gap between intentional and ostensive communication in non-human primates. *Front. Psychol.* **12:718251**

Washburn MF. 1908. *The Animal Mind: A Text-Book of Comparative Psychology*. London: Macmillan

Wheeler BC, Fischer J. 2012. Functionally referential signals: a promising paradigm whose time has passed. *Evol. Anthropol.* 21(5):195–205

Wilson D. 2003. Relevance and lexical pragmatics. *Ital. J. Linguist.* 15:273–92

Wilson D, Carston R. 2007. A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In *Pragmatics*, ed. N Burton-Roberts, pp. 230–59. London: Palgrave

Wilson D, Sperber D. 2012. *Meaning and Relevance*. Cambridge, UK: Cambridge Univ. Press

Zlatev J, Żywiczyński P, Wacewicz S. 2020. Pantomime as the original human-specific communicative system. *J. Lang. Evol.* 5(2):156–74

Zuberbühler K. 2018. Combinatorial capacities in primates. *Curr. Opin. Behav. Sci.* 21:161–69

Zuberbühler K. 2020. Syntax and compositionality in animal communication. *Philos. Trans. R. Soc. B* **375(1789):20190062**

Zuidema W, de Boer B. 2018. The evolution of combinatorial structure in language. *Curr. Opin. Behav. Sci.* 21:138–44

Shows experimentally that combinatorial communication systems are present even in bacteria.

Presents an innovative cognitive analysis of nonhuman primate communication, with conclusions similar to those of this review.

Presents a more detailed review of combinatorial communication systems in nonhumans.

Contents

Errata

An online log of corrections to *Annual Review of Linguistics* articles may be found at
http://www.annualreviews.org/errata/linguistics