

Chapter One:

# What Makes Human Communication Special?

Daniel W. Harris\*

Draft of October 27, 2020

WE ARE SOCIAL PRIMATES—members of a species that arose about two hundred thousand years ago and proceeded to colonize the far corners of the Earth. Our prominent place in the world is owed to our unique psychological characteristics. Among the most important of these is our capacity to communicate—a capacity that enables us to share, accumulate, and store vast quantities of information, that allows us to coordinate our activities in ways that no other creatures can, and that underwrites the bewildering variety of social relations and institutions that structure our lives.

We can appreciate this uniquely human capacity without belittling the other remarkable forms of communication found in nature. Some bacteria engage in quorum sensing: they secrete chemical signals which, when detected in high enough concentrations, alter their conspecifics' gene expression in ways that are adapted to densely populated environments (Miller and Bassler, 2001). Many plants, when attacked by herbivores, emit volatile organic compounds that trigger defensive responses in nearby plants (Heil and Karban, 2010). Pheromones are used for communication by organisms up and down the phylogenetic tree (Wyatt, 2014). For example, Rasmussen et al. (1996) found that a chemical compound used by many insects to attract mates also serves the same function in the urine of ovulating Asian elephants. Honey bees perform a dance whose duration and angle to gravity informs

---

\*This is a draft of the first chapter of a book manuscript that is (tentatively) titled 'Human Communication'. I have received helpful feedback earlier drafts from Josh Armstrong, David Egan, and Heris Stenzel. I welcome feedback, which can be sent to [danielwharris@gmail.com](mailto:danielwharris@gmail.com).

their sisters of the location of distant nectar and pollen (von Frisch, 1927; Lindauer, 1961). Many birds sing complex songs in order to advertise their identities, locations, and degrees of eagerness to mate, or to warn others to keep away (Catchpole and Slater, 2008). Humpback whales coordinate their mating activities with long, looping songs that can travel many miles through the ocean (Whitehead and Rendell, 2015, 76–84). Bottlenose dolphins make signature whistles that uniquely identify the whistler, and can learn to address other pod members by their addressees' signature whistles (Janik et al., 2006; King and Janik, 2013). Prairie dogs use distinct vocalizations to bond, to warn of predators, and to threaten (Waring, 1970). Some primates have been shown to use different alarm calls to warn each other of different categories of predators, as well as other vocalizations for bonding with and threatening one another (Cheney and Seyfarth, 1990, 2007; Schlenker et al., 2016b). The gestures of great apes are among the most sophisticated nonhuman signals, both in how they are learned, in the flexibility of when they are used, and in the range of communicative functions that they serve (Tomasello and Call, 2019). And some animals, including dolphins, parrots, dogs, and apes, have been trained to communicate with human trainers in more sophisticated ways than what they do with conspecifics in the wild.

What sets human communication apart from these phenomena? My aim in this chapter is to give ten answers to this question—to describe ten features of our capacity for communication that make it uniquely powerful and flexible. I should say at the outset, however, that my goal is not to demonstrate human superiority, or even to demonstrate that our capacity for communication is better than the alternatives in every possible way. (nonhumans communicators can also do some things that we can't, after all: If I tried to attract a mate with my underwater singing, or if I tried to dance out directions to the grocery store, the results would be communicative as well as aesthetic failures.)<sup>1</sup> My goal in this chapter is instead to highlight and defamiliarize some features of human communication that stand in need of theoretical explanation, and to do this by pointing out the degree to which they distinguish human communication from comparable phenomena that can be found elsewhere in the natural world.

In later chapters, I will be concerned with the psychological underpinnings of our capacity for communication. In particular, I will be focusing on three psycho-

---

<sup>1</sup>It should go without saying—though I fear it doesn't—that I do not take any of the human capacities that I will discuss in this book to be reasons to think of humans to be more morally significant than nonhuman animals.

logical capacities that come together to make our way of communicating possible: Our capacity to recognize and predict one another's states of mind, our capacity to plan complex actions and coordinate them with others, and our capacity to design complex linguistic evidence of our communicative intentions. The ten features of human communication on which I will focus in this chapter would not be possible if we did not possess highly advanced versions of these capacities. Of the three, I will argue that at least some of the psychological mechanisms underlying human language are entirely absent in other creatures, and those underlying our capacities for planning and mindreading are either unique to humans or so much more powerful in humans than in other animals that they allow for entirely different communicative applications. In short, I take the unique power and flexibility of human communication to be grounded in at least these three psychological discontinuities that lie between us and the phylogenetic tree from which we have sprung.

However, in this chapter I will attempt to describe human communication at a level of abstraction that prescind from the psychological and biological mechanisms that make it work. This will allow me to more effectively compare what we do to what nonhuman communicators do, although they might have very different brain structures and information-processing strategies for doing it. For each feature of human communication that I will single out, I will also discuss what I take to be the closest analogues in other species. The point of these comparisons is not to suggest common implementation details or evolutionary history. (Humans and birds likely don't share a common ancestor from which our syntactic capacities both developed, for example.) Rather the recurring point will be to underscore and map out the gulf between us and the earth's other inhabitants.

Faced with the question of what makes human communication unique, many will be tempted to answer 'language' and leave it at that. Of course, I agree that our capacity for language is an important part of the story, and I will have much to say about it. But there are two reasons not to be satisfied with this answer on its own. First, we should want to know which features, *exactly*, make linguistic communication so different from the sophisticated systems used by animals to communicate. Second, one part of my point in this chapter will be that much of what makes human communication special can also be discerned in the ways that we communicate nonlinguistically.

I will therefore do my best to adopt neutral terminology that abstracts from both the human–animal and linguistic–nonlinguistic distinctions. By the term 'signal', I will mean any behavior that serves as a vehicle for communication, whether

this behavior is the utterance of a sentence, the raising of an eyebrow, a monkey's alarm call, or the involuntary secretion of a chemical. And I will talk about the communicative “point” or “function” of a signal in a way that is neutral about two questions: (i) whether this point is to convey information, to elicit a reflexive behavior, or to adjust a conspecific's gene expression; and (ii) whether a signal has its point as a result of a species-wide biological function, a localized cultural convention, or a sophisticated, one-off plan in the sender's mind. I will normally refer to the ‘sender’, ‘communicator’, or ‘initiator’ of a signal rather than the ‘speaker’, and I will talk about the ‘addressee’ only when the signal is, in some sense, aimed at distinguished recipient. Otherwise I will talk about the ‘receiver’. I will talk of an organism's ‘repertoire’ to describe the collection of signal types that the organism in some sense has stored for potential use.

With these preambles out of the way, I turn now to my list of ten features of human communication that make it unique:

## 1 Our use of structurally complex signals

Human communication can serve an enormous range of functions, conveying more information about more topics in more circumstances than any system used by any other creatures can manage. One reason for this is the enormous number of functionally differentiated signals that we are capable of using. As we will see in later sections, there are several reasons for this limitlessness, all of which arise to some degree in nonlinguistic communication. But I will begin by discussing a feature that is exemplified by, though perhaps not unique to, human language—namely, our capacity to produce complex signals whose communicative functions systematically depend on the details of their structures. It will be important to separate out two aspects of this capacity—our *syntactic* capacity to create and perceive signals with complex structures, on one hand, and our *semantic* capacity to endow these complex signals with structure-dependent meanings, on the other. I will discuss the best analogs of each of these capacities in other creatures. However, the animals capable of using signals with significant syntactic complexity lack any compositional-semantic capacity by means of which to put their complex signals to correspondingly diverse communicative ends, and those animals who may possess some compositional-semantic capacity are severely limited in the syntactic complexity of their signals. It is therefore our ability to combine these two capacities, as much as the unique extent of either capacity on its own, that explains the enormous expressive power of

human language.

First consider songbirds—the most impressive nonhuman syntactic virtuosos.<sup>2</sup> Although some songbirds’ repertoires are made up of discrete song types with few shared elements, others construct songs from a basic repertoire of song elements—either individual notes or in some cases larger, repeatable phrase-like units (Catchpole and Slater, 2008, 204). Even when the resulting songs are built from a relatively small number of elements and have relatively simple structure, a huge number of distinct song types can result. For example, the songs of male sedge warblers are made up of as many as 85 distinct syllables, but performances are sufficiently long and varied in their arrangement of syllables that a single bird may never sing the same sequence more than once in its life (Catchpole, 1976). The songs of some bird species have hierarchical syntactic structure, and admit of syntactic analysis of varying degrees of sophistication. For example, a single “bout” of a zebra finch’s song comprises several introductory notes followed by several repetitions of a “motif” that is made up of a sequence of “syllables”, each of which is in turn made up of an uninterrupted sequence of notes (Berwick et al., 2011, 114). Zebra finch song can thus be analyzed as having four layers of hierarchical structure, though the resulting song structures are highly formulaic. The songs of other species, such as the Bengalese finch and the nightingale, are less formulaic in their structures, but with fewer layers of hierarchical depth (Berwick et al., 2011, 115–118).

Although the structure of birdsong is fascinating, the syntax of human sentences is considerably more complex than this or any other nonhuman communication system, in ways that have been precisely described by linguists (Berwick et al., 2011; Hurford, 2012). Although speech is serial, with sounds and words following one after another, linguists have assembled a large and compelling body of evidence that linguistic expressions have hierarchical structures, at multiple levels of organization. Phonologists study the principles by which words are built out of syllables, and by which syllables are in turn built out of phonemes, which are the smallest linguistic units of sound. (Phonologists also study analogous principles governing the sub-morphological structure of sign language.) Syntacticians study how the most fundamental meaningful units of language—root words and other morphemes, such as prefixes, suffixes, and inflections—are structured to form complex phrases, up to the level of the sentence.

---

<sup>2</sup>There are some other interesting examples. One is the humpback whale, whose songs are typically made up of approximately eight “themes”, each of which consists of as few as two and sometimes more than twenty repetitions of a single “phrase” (Whitehead and Rendell, 2015, 77–79).

Syntactic structures are standardly represented using phrase-structure diagrams, such as the simplified example in Figure 1. This diagram suppresses much of the structural detail that would normally be included by a syntax textbook, including the syntactic manifestations of tense, aspect, clause type, and prosody, as well as various unpronounced expressions that, many would argue, are present in the sentence's underlying structure. But this simplified example gives a sense of how natural-language sentences are structured. In particular, there is a near consensus that each complex linguistic expression is the result of merging two constituents, so that a sentence's structure can be traced from a root node (here, the top S node) down a binary-branching tree to a collection of terminal nodes that represent lexical items—the morphosyntactically primitive units from which more complex expressions are constructed. Even in this highly simplified diagram, we find six layers of hierarchical depth—more than can be found in the songs of Zebra Finches or any other nonhuman signal. Longer sentences may be dramatically more complex.

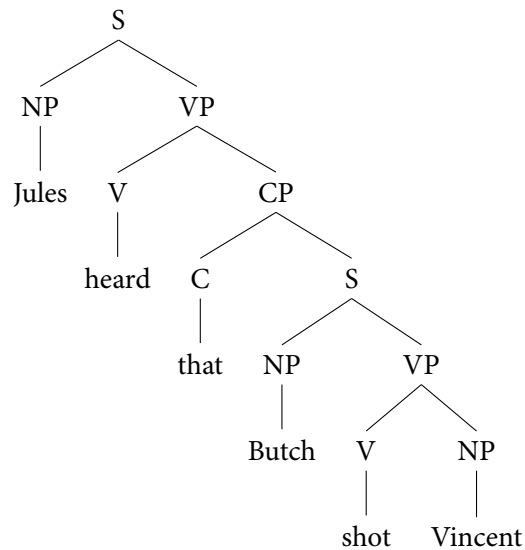


FIGURE 1: A simplified representation of the phrase-structure of the sentence, 'Jules heard that Butch shot Vincent'. Grammatical categories are abbreviated as follows: S = sentence; NP = noun phrase; VP = verb phrase; V = verb; CP = complementizer phrase; C = complementizer.

Figure 1 also illustrates a unique property of human languages, which is that their sentence structures can be recursive. Natural-language syntax allows for ex-

pressions in a given syntactic category to be embedded within expressions of the same category. The sentence in Figure 1 has another sentence, ‘Butch shot Vincent’, embedded within it, for example. The syntactic principles governing natural languages place no principled limits on how many embeddings of this kind can occur, and so recursion makes human languages unbounded in their structural potential. Taking any grammatical sentence, we can append ‘Jules heard that’ to the front to wind up with a new grammatical sentence. Similarly, we can create phrases like, ‘The killer of the killer of the killer of...Marvin’, with the number of intervening possessives as large as we like, without rendering the result ungrammatical. Our ability to use the complex structures that result from recursive embeddings like these is limited only by the time and psychological resources that we are able to invest. By contrast, the evidence suggests that no other organisms use signals with recursive syntax and semantics, or anything approaching the syntactic complexity of relatively simple human sentences.

So, although birds, whales, and some other animals have some syntactic capacities, human languages possess far greater syntactic potential. But this is only one half of what makes human language such a fertile system for communication. The second half is our semantic capacity, which allows us to harness all of this syntactic complexity for communicative purposes by systematically endowing complex expressions with meanings.

To bring this difference into focus, first consider how signal structure and communicative function are related in birdsong. I will use the sedge warbler as my main example, though the point generalizes. Male sedge warblers produce an enormous variety of song structures—probably a new one each time they sing. But this enormous variation in song structure does not correspond to variation in communicative function. Males sing during breeding season for the sole purpose of attracting mates, and stop singing once they have paired with a female (Catchpole, 1973). In a sense, then, the song of the sedge warbler has only one communicative function, which does not vary at all with song structure. This isn’t quite the whole story, however. There is evidence that greater song complexity leads to greater mating success: female sedge warblers show a preference for males with larger repertoires of syllables (Catchpole et al., 1984), and male sedge warblers who incorporate more syllables into their songs tend to pair with mates earlier in the breeding season (Buchanan and Catchpole, 1997; Catchpole, 1980). Similar patterns show up in species for whom song plays a similar role in mating, and the resulting pressure from sexual selection seems to be the primary driver of song complexity in songbirds (Catchpole and

Slater, 2008, ch.7). This explains why some birdsong is so complex, but it involves a much looser connection between structure and function than what we find in human language. For flirtatious songbirds, more complexity is better, but not because greater complexity allows males to say more things to females. Individual song elements (such as syllables) have no specific meanings of their own—they make no particular, isolable contribution to the communicative function of larger songs—and there is no systematic recipe for determining different functions for complex songs on the basis of their different structures and the elements that they include. Although birdsong can serve a variety of functions other than mate attraction, the broad point that communicative function does not systematically depend on song structure is completely general (Berwick et al., 2011).

In this respect, the structure of birdsong resembles the phonological structure of human words rather than the syntactic structure of human sentences. Syntax tells us how sentences decompose into meaningful units—clauses, phrases, and ultimately root words and other morphemes. The units of sound within words that are studied by phonologists, by contrast, are not themselves meaningful, and so there is in general no recipe for deriving the meanings of root words and other morphemes from their phonological parts. Although the word ‘matter’ has its own meaning, neither the syllable /mæt/ nor the phoneme /m/ do. Likewise, although entire bouts of birdsong have meanings (communicative functions), the elements from which they are constructed do not, and there is no recipe for deriving the meaning of a bout from the meanings of its parts. For this reason, Marler (1998) describes the structure of birdsong as a kind of ‘phonological syntax’, and others have adopted the same terminology to describe the structured songs of other animals, including humpback whales and gibbons (Hurford, 2012; Knight, 2001).

Semantic compositionality is implemented by human languages in fascinatingly intricate ways.<sup>3</sup> A semantic theory is a computational model—a recipe—that ex-

---

<sup>3</sup>Even if we ignore natural language, humans are far more prolific and adept creators of complex, semantically compositional signals than any other organisms. One example is our ability to communicate with diagrams. Architectural drawings are highly structured devices of communication, made up of a multitude of nonlinguistic elements and structural relations that encode information about an actual or potential built environment. The weights, colors, and types of lines, a variety of iconographic shapes, and various spatial relations all have particular meanings that contribute to the drawing’s overall content, and that allow architects to communicate novel instructions to contractors. Something similar could be said of maps and the user interfaces of computing devices, which use complex structures of conventional elements to convey information to users. We also use combinations of gestures, emoji, and facial expressions in relatively sophisticated ways. I do not mean to suggest that our use



plains how the meanings of complex expressions are built up from the meanings of their parts in a way that is guided by syntax. A catalogue of lexical entries specifies the meanings of morphemes, and a collection of composition rules enumerates the possible ways of combining the meanings of simpler expressions to yield the meanings of more complex expressions.

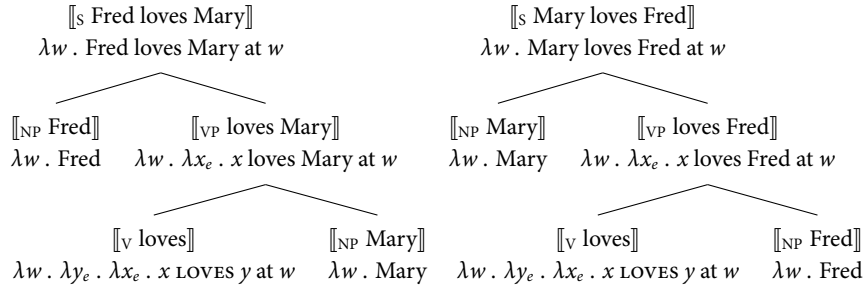
In the most influential theoretical frameworks, meanings are modeled as functions (in the mathematical or computational sense of ‘function’), and composition rules are ways of manipulating and combining these functions in a way that is guided by syntax. The meaning of a declarative sentence is usually thought of as a piece of informational content—a proposition—and is often formally modeled as a function that specifies which possible states of affairs the sentence would be true of (or, equivalently, as a set of possible worlds). For example, the meaning of the sentence ‘Fred loves Mary’ is modeled as a function that maps a possible world to truth if and only if Fred loves Mary at that world. By swapping ‘Fred’ and ‘Mary’, we change their position in the sentence’s phrase structure, which changes the order in which their meanings combine with the verb, leading to a different derivation and a different meaning for the sentence. Figure 2 uses this minimal pair of semantic derivations to illustrate some of the standard technical apparatus and the way that a small syntactic difference can affect the meaning of the resulting sentence. In this way, the meanings of natural-language sentences depend on the meanings of their parts as well as the hierarchical details of their syntactic structure. Because the composition rules of human languages can generate meanings for sentences with highly complex, and even recursive syntactic structures, a compositional-semantic theory is a compact recipe for specifying the meanings of an infinity of phrases and sentences.

As we have seen, the most syntactically complex signals used by nonhuman communicators are not semantically compositional in anything like the foregoing sense. The search for compositionality in nonhuman signals has turned up some possible cases, but there are reasons for skepticism about each case. I will discuss a few.

First consider honeybees, whose dance communicates information about both the distance and direction of a potential food source. These two components of their message are encoded by different aspects of the structure of the dance. Distance cor-

---

of any of these modes of communication displays a form of compositionality that is as rich as the sort of semantic composition involved in language. I will be arguing later in this book that humans have dedicated psychological mechanisms for language, but we don’t have any such mechanisms for architectural drawings, maps, or emoji.



**FIGURE 2:** A minimal pair of simplified semantic derivations in the framework of Heim and Kratzer (1998). We place double brackets around a word or phrase to refer to its semantic value (roughly: its meaning), which is given in lambda notation directly below. For example, the semantic value of the transitive verb ‘loves’ is a function from a possible world  $w$  and an entity  $y$  to a function from an entity  $x$  to a truth value: true if  $x$  loves  $y$  at  $w$ , and false otherwise. In these simplified examples, semantic values compose by a version of functional application (specifically:  $\llbracket \alpha \beta \rrbracket = \lambda w . (\llbracket \alpha \rrbracket(w))(\llbracket \beta \rrbracket(w))$ ), and the semantic value of a sentence is a function that specifies what it takes for the sentence to be true at a possible world.

responds to the time it takes for the bee to complete its waggle run, and direction is indicated by the angle of the waggle run to gravity, which is equal to the angle of the direction of the food source to the sun’s position in the sky at the time of the dance. There is experimental evidence suggesting that bees store and process these two pieces of information independently. Marco and Menzel (2005) forced bees to take an indirect route to a food source, through a tube with a  $90^\circ$  angle, but allowed them to take a direct shortcut back. When they danced upon their return, the length of their waggle runs corresponded to the length of the longer and less direct outbound flight, while the angle of their dance corresponded to the direction of their shorter and more direct inbound flight. This suggests that we might want to think of the waggle dance as a compositional system in which distinct elements (duration and direction) encode distinct information (distance and direction) which are systematically integrated in the performance of the dance. However, we might wonder whether the bees really have to integrate the two pieces of information in anything like the sense that is relevant to compositional semantics. Perhaps the sender stores the two pieces of information separately, using one to determine which direction to face and the other to determine how long to dance, and the receiver likewise reads off these two pieces of information separately from the dance and then uses one to

determine which direction to fly in and the other to determine how long to fly. It is unclear whether we need to think of bees as semantically composing the two pieces of information, in any interesting sense, in order to explain how they communicate.

A second possible case of semantic composition concerns some species of ants, who can instruct each other how to navigate through mazes by means of a system that involves rubbing their antennae together. This system has not been decoded, and so we don't know how ants package information in antenna-rubs, but some of its information-theoretic properties have been studied (Reznikova, 2007; Reznikova and Ryabko, 1986). Instructions normally take longer for more complex mazes, but not when the instructions lend themselves to certain forms of compression: ants are quicker to communicate their equivalent of 'take six lefts' than their equivalent of 'left, right, right, left, right, left', for example. A possible explanation of this sort of information compression is that some form of semantic composition may be at play. However, at present this explanation is rather speculative (Hurford, 2012, 11–12).

What about cases that are closer to humans? Some of the most interesting recent work has focused on possible semantic composition in monkey communication. Putty-nosed monkeys use one alarm call, 'pyow', to sound a general alert, a second call, 'hack', to warn of aerial predators, and sequences of pyows followed by hacks to signal, non-urgently, that the group should move. Most theorists have concluded that the meanings of pyow-hack sequences are not systematically composed from the meanings of their parts, but are akin to human idioms, such as 'see the light' in English, whose meanings bear at most historical relationships to the meanings of their parts (Arnold and Zuberbühler 2012; Hurford 2012, 14–15). However, Schlenker et al. (2016a,b) have recently argued that a compositional analysis of 'pyow-hack' sequences is consistent with the data, provided that we also assume that putty-nosed monkeys obey certain pragmatic principles, such as the principle that more urgent portions of a sequence of calls always precede less urgent sequences.

A second potential example of monkey compositionality involves the alarm calls of Campbell's monkeys, who utter the signal 'krak' to warn of leopards (or, in some groups, as a general warning), and 'hok' to warn of serious aerial threats, such as eagles. But Campbell's monkeys sometimes also append a kind of suffix, '-oo', to either of these alarm calls, and in each case it serves to communicate less urgency than the root call alone would indicate (Ouattara et al., 2009). Because of this minimal form of systematicity, Schlenker et al. (2014, 2016b) have speculated that a kind of semantic composition is at work, though others have denied that this is the best explanation of the data (Fitch, 2016; Hurford, 2012).

Perhaps the most obvious reason to doubt that the calls of either species of monkeys are semantically compositional is that this hypothesis just isn't needed to explain their behavior. In humans, there is no plausible alternative explanation of the fact that we can use and understand an unbounded number of complex expressions, including those with entirely novel sentences. By contrast, the alleged cases of semantic composition by monkeys involves increases of only one or two signals—something that can be explained by supposing that the complex signals are semantically primitive but just happen to use the same vocalizations. Schlenker et al. (2016b) argue that the compositional analyses may be more parsimonious, but in the end the question will come down to as-yet-undiscovered facts about monkey cognition.

A fourth source of potential examples of semantic composition in animals comes from studies of animals held in captivity who have been trained to communicate with human-language-like signals. The most publicized examples involve great apes.<sup>4</sup> Although great-ape gestures and vocalizations are sophisticated in ways that I will discuss in later sections, one thing that our fellow great apes don't do is to communicate with each other using complex signals whose meanings are derived from the meanings of their parts (Tomasello, 2008, 30). There is, however, a much-publicized tradition of attempting to teach systems of lexigrams or sign-like gestures to apes who are being held in captivity. The most famous such apes are the chimpanzees, Neam ("Nim") Chimpsky (Terrace et al., 1979) and Washoe (Gardner and Gardner, 1969), the bonobo Kanzi (Savage-Rumbaugh and Lewin, 1994), and the gorilla Koko (Patterson, 1978; Patterson and Matevia, 2001). The researchers who have carried out these experiments have sometimes made bold claims about the linguistic abilities of the apes they've trained. For example, nearly three decades after they began training Koko, Patterson and Matevia (2001, 167) reported that she "combines her working vocabulary of over 500 signs into statements averaging three to six signs in length", and that "her emitted vocabulary—those signs she has used correctly on one or more occasions—is over 1,000." They also reported that Koko "understands spoken English, and often carries on 'bilingual' conversations, responding in Sign to questions asked in English" (Patterson and Matevia, 2001, 165). Greenfield and Savage-Rumbaugh (1990) report that after training Kanzi to communicate with lexigrams, Kanzi began to use "spontaneous combinatorial two-

---

<sup>4</sup>But see also Kako (1999) for a survey of case studies on African gray parrots and bottlenose dolphins, and Pilley (2013) for a discussion of a border collie who responds differently to grammatically different commands, such as 'to ball take frisbee' and 'to frisbee take ball'.

symbol utterances” consisting of multiple lexigrams or lexigram–gesture combinations. Greenfield and Savage Rumbaugh assembled and analyzed a corpus of 723 such utterances produced by Kanzi over the course of a single week, and argue that many of these combinations constitute meaningful grammatical combinations. A claim that is often made about these apes is that their linguistic abilities are roughly equivalent to those of young human children—a claim that is thought to support the view that there is an evolutionary link between our capacity for language and those of the other great apes, as phylogenetically related traits tend to be most similar early in ontogenetic development (Greenfield and Savage-Rumbaugh, 1990). If these claims are borne out, then they would support the surprising conclusion that non-human great apes have a latent linguistic capacity that goes wholly unused in the wild, but that can be brought out by human trainers.

Critics of ape sign-language research have argued that they rest on superficial analogies between ape behavior and human language, and that apes’ supposed multi-symbol utterances show no evidence of genuine syntax or compositional semantics: they are disorganized sequences, or they are learned from trainers as idiomatic units, or they are random meaningful-sounding examples plucked from a sea of gibberish.<sup>5</sup> The most important criticisms of ape signing research have been grounded in worries about bias in the gathering and coding of data, compounded by the fact that the evidence is gathered by the apes’ own trainers, who often have intimate relationships with their subjects. Since the process of gathering data has typically not been filmed and made available to other researchers, it usually cannot be independently checked. These worries have tended to be borne out on the few occasions when ape researchers have released video evidence of their experiments. For example, analysis of video evidence revealed that around half of Nim’s signs consisted of direct imitations of humans with whom he was interacting—something that had not been previously noticed (Terrace et al., 1979). Recent, more sophisticated analyses of video corpora of chimpanzee signs have suggested that ape signing lacks genuine syntax or compositional semantics, and that analogies to the grammatical capacities of human children are unsupported (Rivas, 2005; Yang, 2013).

So, do bees, ants, monkeys, or captive great apes use syntactically complex, semantically compositional signals? I doubt it, but I have not tried to settle the issue here. My point in this section ultimately doesn’t hang on this question. After all: even if all of the supposed examples of nonhuman compositionality that I’ve just described are real, what they show is that some non-human animals have the ability

---

<sup>5</sup>See, for example, Seidenberg and Pettito (1979); Wallman (1992); Pinker (1994, 334–342).

to use semantically compositional signals that are far less syntactically complex, and so far less information rich, than even the simplest human sentences.

## 2 Our large signal repertoires

Grammatical complexity aside, we construct our sentences from enormous collections of building blocks. Linguists estimate that normal monolingual adults have between 50,000 and 100,000 distinct items in their vocabularies, and these figures may be much larger for multilingual adults (Clark, 2009, 75–76). Each of these vocabulary items is a morpheme—a semantic primitive whose meaning cannot be inferred from those of more basic symbols. If we measure humans’ repertoires of sentences, their size becomes effectively unlimited.

I know of no systematic attempt to count the nonlinguistic signal types in our repertoires, but there are good reasons to think that they too must be many. Consider smiling, which it may be tempting to think of as a single signal type with a single meaning. This turns out to be wrong. Ekman (1987) tells us that “there are dozens of smiles, each differing in appearance and in the message expressed” (150). He goes on to give eighteen examples of distinct smiles in the repertoires of all of the subjects he has studied, distinguishing their outward appearance, the specific facial-muscle contractions involved in producing them, and their characteristic communicative functions (150–161). This multifariousness may have been what Herman Melville (1852) intuited when he had one of his protagonists say that “a smile is the chosen vehicle for all ambiguities”. And of course, our nonverbal signaling repertoires include an enormous number of other facial expressions, gestures, traffic signs, user-interface elements, icons, logos, emoji, pictures, and so on.

Most nonhuman communicators have comparatively tiny repertoires of primitive signals or signal-elements. Nonhuman primates possess repertoires that number in the single or low double digits, including both vocalizations and gestures. Some great apes have been trained by humans to have larger repertoires. For example, apes do not point in the wild, but sometimes learn to point in captivity as a way of requesting things from humans (Tomasello, 2008, 34–38). And apes who have undergone intensive training may have learned to produce and respond to hundreds of lexigrams or sign-like gestures (Patterson and Matevia, 2001). Similarly, some dogs have been trained to respond to hundreds of distinct verbal commands that resemble human words (Kaminski et al., 2004). In the most extreme documented case, a border collie named Chaser was taught to respond to distinct names for over

1000 toys (Pilley and Reid, 2011).

The animals with the largest repertoires of signal types are songbirds. Among birds who creatively build their songs from repeatable elements, repertoires may include dozens of elements. For example, a sedge warbler constructs its countless song types from about 80 syllables (Buchanan and Catchpole, 1997). Other birds have repertoires of dozens or even hundreds of distinct, stereotyped song types that share few elements (Catchpole and Slater, 2008, 205). An extreme case is the brown thrasher, a relative of the mockingbird that inhabits eastern North America, and whose repertoires have been estimated to include between 1400 and 2300 song types (Kroodsma and Parker, 1977). This dazzling variety of birdsongs does not correspond to functional variety, however, since distinct song types generally don't have distinct communicative functions. This is a crucial difference from humans when it comes to explaining our communicative capacities.

### 3 The extensibility of our signal repertoires

Part of what explains the immensity of our signaling repertoires is our ability to add new signal types to our repertoires. Although we sometimes learn the meanings of new signals through explicit instruction, in practice we often work out the meanings of novel signals on the fly.

Our vocabularies are not just big; they also vary considerably between speakers of the same language. You likely know many specialized words that are connected to your profession and hobbies, for example. This raises the problem of how we manage to cope with unfamiliar words when we encounter them. We could seek an explicit definition, but there is ample evidence that this is not how humans learn the majority of the words that we know. Consider the task faced by a normal two year old, whose vocabulary will swell from a few hundred items to around 14,000 by the time they are six. This means that they will have to acquire a new word for every hour and a half that they are awake over the next four years (Pinker, 1994, 151). Most of this learning happens without explicit instruction, with children inferring the meaning of a new word on one of the first few times they encounter it—a phenomenon that psychologists call “fast mapping” (Carey and Bartlett 1978; Bloom 2000, ch.2). We retain our fast-mapping ability into adulthood, although it drops off significantly in adolescence.

The ability to work out the meaning of a novel signal without having it explained using signals already in our repertoires is not limited to language learning. Scott-

Phillips et al. (2009) have given compelling experimental evidence of this fact. They devised a computer game for two players in which each player moves a figure around a four-by-four grid of randomly arrayed colors. In each successive round, the colors are reshuffled, and the players' aim is to move their figure to the same color. But although each player can see the position of the other's figure in their grid, neither can see how the colors are arrayed on the other's grid. The players cannot see or hear each other; the only way that they can interact is to move their figures around the grid as much as they want before settling on a square. But there is no pre-existing system by means of which they can use this information channel to communicate. The finding of Scott-Phillips and his colleagues was that a communication system spontaneously emerged more than half the time when repeated rounds of this game were played. Some pairs developed systems in which different patterns of movement were tied to different messages: a C-shape meant that they should go to yellow, an up-and-down oscillation meant green, and so on. By the end of this process, they had worked out a novel system of signals, but they did so without any prior system in which to do the working out.

Many nonhuman communication systems are not learned at all: they are universal to the normal members of a given species, and show no variation. There are some interesting exceptions. Honeybees' waggle dance is almost environmentally invariant, though not quite completely. In order to correctly interpret the dance, bees must be aware of the current location of the sun, and it has been shown that they are capable of calculating its position even when it is not visible—a latitude-specific trick. Lindauer (1961) reports that bees who are moved from the northern hemisphere to the southern hemisphere (or vice versa) consistently misinterpret the dance for the rest of their lives, but their offspring do not make the same mistake, suggesting that this aspect of bees' ability to decode the dance is a skill that is acquired during a critical period during bees' cognitive development and thereafter fixed.

Most monkeys and apes who use alarm call have a universal, species-specific repertoire of calls (Tomasello, 2008, 16–17), though there are some documented instances of minor variations in the meanings of some of these calls across populations (Schlenker et al., 2014). Many songbirds learn to sing from older 'tutors,' exhibit dialectal variation between populations within a species, and innovate idiosyncratic song elements (Catchpole and Slater, 2008, ch.3). Humpback whales sing songs that vary between pods in different geographic regions and evolve over time. Noad et al. (2000) discovered that a sudden revolution in the type of song sung



by eastern-Australian humpbacks between 1996 and 1998 was due to the influence of a small number of wayward western-Australian humpbacks whose dialect caught on over the course of two years—an event that has been compared to the effect of the British invasion on 1960s American popular music (Jensen, 2000). However, any given bird or whale is limited to a small repertoire of song at any given time, and new song elements tend to push out old variations. There is, moreover, little evidence that variations in song elements and structure systematically correspond to differences in what is communicated, and some biologists have argued that much variation may be due to random drift rather than the enhancement of communicative function (Slater, 1986). None of these forms of learning lead to anything like the explosion in communicative potential of the kind exhibited by human children.

Some of the most impressive possible examples of nonhuman signal learning have involved animals who have been trained by humans to produce or understand signals that resemble elements of human language. According to her trainers, Koko the gorilla learned to use hundreds of sign-like gestures with distinct meanings, and could understand over 2000 words of spoken English (Patterson, 1978; Patterson and Matevia, 2001). A parrot named Alex was trained to communicate with over 100 sounds that resemble English words (Pepperberg, 1999). Border collies have been trained to respond differentially to as many as 1000 distinct labels for toys by fetching them on command, and have even managed something like fast mapping: given an unfamiliar command in the presence of an array of toys, all but one of which has a familiar name, border collies are able to infer that the new command labeled the unfamiliar object (Kaminski et al., 2004; Pilley and Reid, 2011). All of these studies have generated controversy, both on methodological grounds and because of disagreement about what, if anything, they tell us about the evolutionary and developmental origins of human language. For present purposes, we can just note that even if we were to take these studies' boldest findings at face value, we would conclude only that the animals in question have a capacity to acquire a collection of signal types that is two or more orders of magnitude smaller than the repertoire of a normal human child, and that they never actually manifest this capacity in their natural environments.

## **4 Our use of novel signals**

The ease with which humans add to their signaling repertoire on the fly is related to another distinctive human characteristic—namely, that we are able to communicate

using novel signals that have no pre-established meanings for either the communicator or the recipient. You might do your roommate's chores in order to passive-aggressively communicate your annoyance at their messiness. Or you could play a particular song at the right moment in order to communicate that it's time to dance. You can communicate by honking your car's horn, by slamming a book down on a co-worker's desk, by looking your sibling straight in the eye while flushing their goldfish down the toilet, by suddenly turning your car around after your child has once again kicked the back of your seat, or by placing the severed head of someone's horse in their bed while they sleep. One reason why there is no upper bound on the number of signals that we can use to communicate is that there seem to be no restrictions on the kinds of behaviors that can be spontaneously turned into signals with a discernible communicative point, given the right circumstances.

Every instance of on-the-fly word learning is also an instance of communication via a signal-type that is meaningless for the word-learner prior to successful communication. In situations like these, one has to work out what the speaker means in order to infer the meaning of the word they've uttered and add it to one's repertoire. Border collies' ability to infer the referent of a novel name for a novel toy thus gives us one of our few clear examples of nonhumans communicating with novel signals. But of course, this ability is limited in ways that human use of novel signals is not: border collies only do this when communicating with humans, only when they are on the receiving end, and only when the novel signal is a name for a toy.

More extreme cases of novel signal use occur when people with radically non-overlapping repertoires of signals come into contact. Most of us have encountered minor versions of this phenomenon when trying to order at a restaurant in a foreign country. Communicating in these scenarios is difficult, but it can usually be done. The most impressive examples have occurred during first encounters between hitherto-isolated populations of humans with no shared culture or social conventions. Studies of these situations show that humans tend to fall back on a mixture of cross-cultural universals and signals whose meanings can be inferred by unconventional means (Levinson, 2006). For example, studies of encounters between users of different sign languages have shown that they manage to communicate by exploiting iconic properties of signs that can be recognized independently of the rules of any particular sign language (Byun et al., 2018). And of course, there is a sense in which the most extreme first-contact scenario was not that of James Cook or Christopher Columbus but rather the one that each of us faced as a prelinguistic and uncultured child.

One could try to explain away all of these examples by positing an innate signaling repertoire shared by all humans, on which we fall back when our richer learned repertoires fail us (Levinson, 2006). This is part of what makes the computer-game experiment designed by Scott-Phillips and his colleagues, which I discussed in Section 3, so ingenious, since it artificially blocks all of the usual information channels that our innate signaling methods might exploit. At early stages of the game, participants are somehow able to make their proposals for new signals understood, despite having no pre-existing repertoire of signals to fall back on.

## 5 Our ability to communicate indirectly

Even in situations when human communicators possess a shared repertoire of meaningful signals, their use of these signals is invariably much more flexible than what we can find among other species. Consider the fact that we routinely convey information that outstrips or replaces the usual, literal meanings of the sentences that we utter. Suppose, for example, that you're on a date that isn't going particularly well. After an awkward silence, your date says,

(1) I really need to wake up early tomorrow.

It may be true that your date needs to wake up early. But you correctly understand that an additional part of your date's point is that they want to leave. This is no part of (1)'s literal meaning. After all, (1) could be uttered in a different context with the opposite added message. Suppose you are working late with a colleague to finish a project, and you float the possibility of taking a long dinner break before returning to the project. If your colleague were to reply with (1) in this context, their point would likely be that they want to stay until the job is done so that they can get to bed. Finally, imagine a scenario in which you are out with a friend who is notorious for their nocturnal partying lifestyle. Both of you are invited to an after-hours club. You ask your friend if they want to go. With a mischievous raise of their eyebrows, they utter (1). In this scenario, your friend is being sarcastic, and the right thing to do is to interpret them as meaning the opposite of what their sentence indicates.

This point is not restricted to linguistic communication. Consider the metaphorical use of emoji. According to the website *emojipedia*, which conducts research into the use of emoji, only about 7% of uses of the peach emoji on twitter are used to talk about the fruit. Almost as often, it is used to indicate that the user is feeling peachy. By far the most common use is anatomical, to symbolize human buttocks

(Azhar, 2016). The eggplant emoji likewise commonly serves as a symbolic penis, and many other emoji have similar extended uses. Of course, many of these uses have now become conventionalized, but these uses began their lives as metaphors that exploited the iconic connotations of the emoji to guide the reader to the correct interpretation. This is just one nonlinguistic instance of our general capacity to creatively go beyond signal types' standard meanings. All facial expressions, gestures, and other nonlinguistic signal types can similarly be used ironically, indirectly, and metaphorically, given contexts in which addressees are able to infer what the communicator has in mind.

The only nonhuman phenomena that even vaguely resemble this kind of indirect communication are great apes' use of attention-getting gestures (Tomasello, 2008, 27–30). For example, a juvenile chimpanzee might slap the ground in order to draw another's attention to the fact that they are making a playful face with the goal of initiating play (Tomasello and Call, 2019, 463). Tomasello describes this form of communication as “indirect” because it is aimed at achieving one communicative goal (initiating play) by means of another (getting attention). Because the distal goal can vary independently of goal of getting attention, attention-getting signals are more flexible in their communicative uses than perhaps any other kind of nonhuman communication. However, this flexibility is still severely limited as compared to human indirect communication: the proximate communicative goal can only be to draw attention to something in the immediate environment, and the distal goal can only be to elicit some behavior that would normally result from the addressee's observing whatever their attention is being drawn to.

By contrast, human indirect communication can be used to communicate anything that can be communicated directly. Indeed, there are some reasons to think that we can communicate some things *only* indirectly. One reason is that literal paraphrases of metaphorical speech typically leaves out some of what the metaphor conveys—possibly because metaphors function to inspire particularly rich, perspectival representations in the minds of their audiences (Camp, 2006). Indirectly communicating something is a common first step to introducing a new signal into our repertoires by means of which we can later communicate the same thing literally. I have already illustrated this phenomenon with the peach emoji, but it also goes for verbal dead metaphors, such as ‘fall in love’ or, for that matter, ‘dead metaphor’.

## 6 Our use of background information in interpreting signals

Even when we use signals directly and in conformity with pre-established literal meanings, these meanings often fail to fully specify our communicative aims, and so our addressees need more information than they can extract from the signal itself in order for communication to succeed. I will refer to this phenomenon as semantic underspecification.<sup>6</sup> A mild example can be seen in (1)—‘I really need to wake up early tomorrow’—which could be used to communicate different information in different circumstances. Even when a speaker is using (1) in a completely literal and direct way, their addressee needs extralinguistic information, beyond what can be decoded from the sentence itself, in order to understand what has been said. To see what I mean, imagine that you’ve just gotten a new phone and haven’t yet transferred your contacts into it. Having earlier invited several friends over, you receive (1) as a text message, but don’t know from whom. You can infer that *one of* your friends can’t make it, but you can’t tell which one—a piece of information that you would need in order to fully understand their message.

Much nonhuman communication does not involve any underspecification at all, but functions as a simple code: the sender sends a signal when they are in a certain state, and reception of this signal puts the receiver into the state that it is the function of the signal to produce.<sup>7</sup> Pheromonal communication works this way, as does bacterial quorum sensing, although the latter requires the same signal to be sent by many senders before the relevant effect can be produced.

But there are some interesting examples of underspecification in animal communication. Honeybees need information about the current location of the sun in order to correctly act on their sisters’ waggle dances, and they are capable of combining the information that they get from the dance with information about the sun’s location in surprising ways. Bees who are detained in their hive for several hours after witnessing a dance will compensate for the sun’s intervening movement when setting off to follow the earlier instructions (Lindauer, 1961). And bees who

---

<sup>6</sup>My term is borrowed from Neale (2016), who only considers the linguistic case, and of whose notion mine is a broadening. Another term commonly used for this phenomenon in the linguistic case is ‘semantic underdetermination’ (Carston, 2002; Sperber and Wilson, 1995).

<sup>7</sup>Of course, *some* human communication works this way too. Humans are widely thought to use pheromones, for example, though the details of how are not yet as clear as they are in animals (Wyatt, 2014, ch.13). And there may be some information encoded in our speech that does not require supplementation by the addressee.

have only ever been allowed to see the sun during the afternoon will, when released on an overcast day in the morning, correctly perform the dance with respect to the sun's morning position, although they cannot see it and have never seen its position at this time of day (Dyer and Dickenson, 1994).

A range of studies show that primates bring to bear contextual information about the identity of the signaler, their position within the group's social structure, and their credibility when interpreting vocal signals (Armstrong, 2019; Seyfarth and Cheney, 2015). For example, vervet monkeys selectively ignore alarm calls from a group member who has issued too many false alarms of the same kind in the past (Cheney and Seyfarth, 1988). Many primate signals function to threaten or bond with particular other group members, and primates are adept at understanding who the addressee of such signals are, since this is crucial to determining their broader social meaning (Seyfarth and Cheney, 2015). Female baboons are less likely to react with fear to threat-grunts from more dominant females with whom they have recently groomed, for example, because this is a cue that they are not the one being addressed (Engh et al., 2006). Perhaps my favorite example of the broader social context affecting how primates interpret signals is something that will seem uncanonically familiar to any human parent who has taken their toddler to a neighborhood playground: when female baboons hear vocalizations indicating that a young group member is involved in an altercation, they look not at the source of the vocalization but at the young baboon's mother (Cheney and Seyfarth, 1999). All of these phenomena can be thought of as involving semantic underspecification because they make the successful functioning of a signal dependent on the addressee's possession and deployment of information that cannot be gleaned from the signal itself.

Even the socially-aware signal interpretation of nonhuman primates pales in comparison to the kind of semantic underspecification found in human communication, however. Consider the linguistic case first. Obvious examples of semantically underspecified words include certain uses of personal pronouns, tensed expressions, perspectival words that refer to times and places, such as 'here', 'there', 'now', 'tomorrow', and 'today', and demonstrative pronouns like 'this' and 'that'. Over the last several decades, however, philosophers and linguists have documented far more underspecification in human language than seems obvious at first glance. Consider this mundane example:

- (2) Everyone should be tanned before the party.

This sentence could be used to literally and directly communicate many different

things on different occasions. By ‘everyone’, a speaker presumably would not mean everyone in the world, but rather some more restricted group of people (Barwise and Perry, 1983; von Stechow, 1994). But which group in particular? The modal auxiliary, ‘should’, can be used to make a normative claim about what would be best for some purpose or an epistemic claim about what is likely given some background information (Kratzer, 1981). Which kind of claim is being made here? If ‘should’ is being used in a normative way, the speaker may be either ordering their addressees to get tanned or merely reporting rules made up by someone else (Kaufmann, 2012). What kind of speech act is the speaker performing on this occasion? ‘Tanned’ is a gradable adjective, and so to say that someone is tanned is to claim that they are more tanned than some contextually operative standard requires (Kennedy, 2007). How tanned must one be in order to meet the threshold in this case? There are many parties in the world, but by uttering the definite description, ‘the party’, the speaker presupposes that there is some unique party that their addressee will take them to be singling out (Heim, 1983; Neale, 1990). Which party is being singled out in this case? You haven’t understood a speaker who utters (2) unless you have correct answers to all of these questions. Arriving at correct answers requires integrating a wide range of nonlinguistic background information with the information linguistically encoded in the sentence. This is something that we do every time we hear someone speak, often quickly and effortlessly and without noticing that we’re doing it.

Most of the nonlinguistic signals in our repertoires are even more semantically underspecified than the sentences that we utter. Consider the use of a car’s horn. There are *some* ways to differentiate your communicative goal by varying a honk’s properties: shorter honks normally come off as more polite and longer seem more aggressive and exasperated. But on its own, a honk can give those who perceive it little more information than that the honker intends to communicate *something* to *someone* who is nearby. The addressee and intended message must normally be inferred from information found elsewhere in the context, and there are few fixed rules about where to seek it. This extreme form of underspecification undoubtedly limits the usefulness of car horns as communicative devices (a pet peeve of mine!), but we do sometimes manage to successfully communicate by honking. Something similar can be said of most other nonverbal signals. Pointing, making eye contact, raising one’s eyebrows, and most forms of gesture can be used to mean a range of things on different occasions, and correctly interpreting any of these signals may require considerable background information.

## 7 Our ability to perform communicative acts of different kinds

We do more than inform each other of things when we communicate; we also make requests and issue commands, make promises and threats, ask questions, greet, thank, apologize, and congratulate one another, and more. Because these actions have normally been studied by philosophers of language, they are often known as ‘speech acts’—or, more specifically, ‘illocutionary acts’ (1962) or ‘communicative acts’ (1979)—and are said to differ with respect to their illocutionary force.<sup>8</sup> And, indeed, language gives us convenient ways to clearly signal what sort of communicative act we are performing. Every natural language includes a small number of clause types, including declarative, imperative, interrogative, and exclamative, each of which is stereotypically used to perform a distinctive kind of communicative act (Portner, 2004). For example, declarative clauses, such as ‘it’s going to rain this afternoon,’ are normally used to perform assertoric acts, such as statements or predictions, while imperative clauses, such as ‘buy me a drink,’ are normally used to perform directive acts, such as requests or commands. We can also often make the sort of communicative act that we are performing clear by using an explicit performative sentence, such as ‘I hereby promise to feed your chickens’ (Austin, 1962).

Many of the same distinctions can also be discerned in our nonlinguistic communication. Visiting a bakery in Finland, where I do not speak the language, I might request a certain pastry by pointing at it, smiling at the clerk, and raising one finger. And I could then inform them that I enjoyed it by smiling and rubbing my belly. Were I speaker of Finnish, I could have accomplished roughly the same things by uttering translations of the imperative, ‘give me one Rönttönen,’ and then the declarative, ‘that was delicious.’ Language allows me to make the force of my act transparent, but various combinations of context and nonlinguistic convention, together with my addressee’s assumptions about my goals, are often sufficient to do the same.

A notable feature of human communication is that the same signal-types can

---

<sup>8</sup>Austin distinguished illocutionary acts from mere acts of uttering meaningful words (which he called ‘locutionary acts’) and acts of influencing others in various ways, such as the act of offending someone (which he called ‘perlocutionary acts’). Bach and Harnish (1979) further distinguished *communicative* illocutionary acts from *conventional* illocutionary acts, which differ in that the latter can be performed only the context and jurisdiction of often-ritualized social conventions. I will defend this distinction in a later chapter. My focus here is on communicative illocutionary acts (or just ‘communicative acts’).



often be used to perform communicative acts with different illocutionary force. On one occasion, pointing may serve as a way of requesting a pastry; on another, it could serve as a way of informing someone about the presence of an exotic bird that has just landed nearby. And although each clause type is stereotypically associated with a certain kind of illocutionary force, these associations admit of many exceptions. Interrogatives are normally used to ask questions, but the interrogative, ‘could you pass the salt,’ is normally used to both ask a question and request salt. Declaratives are stereotypically associated with assertions, but uttering ‘you’re standing on my foot’ could also be either a request for someone to move or a way of flirting, depending on the circumstances. Declarative sentences that include deontic modals, like ‘you mustn’t park your car there,’ can be used literally to either describe a consequence of some rules or to direct the addressee to do something (Kaufmann, 2012). More generally, it is worth noting that different clause types differ in their syntactic properties, but not in the morphological units from which they are constructed. Nearly any word of any natural language can feature in utterances used to perform communicative acts with all kinds of force.

It is controversial whether it makes sense to distinguish communicative acts with different force among the signals of nonhuman communicators. Of course, there are some tempting cases. For example, Silk (2001) argues that the affiliative and aggressive signals of nonhuman primates should be understood by analogy to promises and threats, respectively. As a case study, she considers the “grunts” and “girney” used by rhesus macaques to signal their benign intentions toward one another. After surveying a range of data—for example, that higher-ranking female rhesus macaques who make these sounds when approaching lower-ranking females are 30 times less likely to initiate aggressive interactions (Silk et al., 2001)—Silk concludes that “these kinds of calls function as pledges or promises to behave peacefully, even when it might be in their short-term interest to behave aggressively” (Silk, 2001). However, this example is problematic because of the difficulty of distinguishing a promise from an act of offering information about one’s intentions or dispositions. Would be better to translate the rhesus macaque’s grunts and girneys into English as ‘I promise not to harm you’ or as ‘I am currently not angry at you’? In the human case, this distinction is usually drawn by saying that genuine promises give rise to “promissory obligations”—a kind of normative status that makes promises unethical to break (Habib, 2018). But it is far from clear that rhesus macaques or other nonhuman primates can partake in obligations of this kind, and this makes the distinction between promising and merely giving information difficult to locate.

In part because of this sort of difficulty with drawing fine-grained distinctions, the most commonly studied case is the distinction between informative, assertion-like acts and directive, command-like acts. However, even in the case of this relatively rough-grained case, it is difficult to get a grip on the distinction in nonhuman communicators. Consider the question of whether the honey bees' dance functions to give information about the location of honey, or to direct other bees to go to that location. The same problem arises for the alarm calls of birds, prairie dogs, and primates: should we translate them as "there is danger nearby" or as "take cover"? Likewise, the mating calls of birds and whales could be understood as meaning "I am over here and eager to mate" or "come and mate with me". Many nonhuman signals apparently have this combined, informative and motivational role. Millikan (1995) has dubbed representations with this dual function "pushmi-pullyu representations", and has argued that relatively unsophisticated representational systems contain only representations of this kind (see also Papineau 2001). This may include many or all of the systems by means of which nonhuman organisms communicate (Harms, 2004a,b).

Not all theorists have agreed. For example, Zollman (Zollman, 2011) considers Zuberbühler's (2001) observation that Campbell's monkeys' response to an eagle-specific alarm call is context dependent: depending on a hearer's location and their information about the location of the eagle, they engage in different escape behaviors. Zollman argues that this gives us a reason to think of the alarm call as being more like an assertion than a directive: it gives the recipient information about the predator, which they then have to decide how to use. Although this is the sort of evidence we would need, however, it is far from decisive, since it doesn't rule out the possibility that the alarm call simply has a relatively unspecific directive meaning, akin to 'protect yourself from the eagle!', which recipients have to implement in different ways depending on context. Compare this with my example of informing a Finnish baker that I enjoyed their pastry—a signal which does not have the function of eliciting any action on the part of the baker, however unspecifically that behavior is described.

There is likewise little reason to think that nonhuman animals ever use the same signal type to perform communicative acts with different force. And there is some evidence that even other great apes are not capable of acquiring signals with this kind of ambiguity. Consider pointing, which humans routinely use in either assertoric or directive ways, as per my example above. Although great apes do not use pointing in order to communicate with one another, they often learn to point in cap-

tivity as a way of communicating with humans. However, apes only ever point in order to request things, and do not point simply to draw attention or to inform, even in situations that regularly elicit informative pointing from human infants (Tomasello, 2008, 34–38). Similarly, apes who have been taught sign-language-like gestures use them almost exclusively in order to request the referent of the sign, and there is meager evidence of clearly informative uses (Rivas 2005; Tomasello 2008, 38).

There is of course room for much more evidence gathering here. And some of the difficulty of applying the concept of illocutionary force to nonhuman communication may be due to theoretical unclarity. In practice, the theory of communicative acts tends to begin with apparently clear, pre-theoretic distinctions that turn out to be quite difficult to theorize in practice.<sup>9</sup> But it is relatively safe to say that human communication is replete with illocutionary distinctions that are either rare or nonexistent among other creatures.

## **8 Our flexible decisions about when, what, and with whom to communicate**

The use of semantically underspecified signals requires flexible sensitivity to background information on the part of receivers. We also display considerable flexibility in the kinds of situations in which we decide to send signals at all, and in the degree to which the communicative functions of our signals may be detached from the situations in which we send them.

For the most part, nonhuman organisms signal in a limited range of situations, and what they communicate is usually closely tied to predictable features of the environment that they are in. Although it is notable that honeybees dance to communicate information about distant resources that their audiences have never experienced, they do this only soon after they have returned from these resources, and only with the function of sending their hive-mates there. Many primates will issue alarm calls only if they take their group members to be in earshot, indicating that they are sensitive to whether they have an audience (Cheney and Seyfarth, 2018; Seyfarth and Cheney, 2018). And some primates continue to issue alarm calls until all of their group members do the same—a mechanism for ensuring that their addressees have understood (Wich and de Vries, 2006). However, primates produce alarm calls only as a direct result of perceiving threats (Cheney and Seyfarth, 1990).

---

<sup>9</sup>For a survey of recent theoretical disagreements, see Harris et al. (2018).

Primates' signals for affiliation and aggression are produced in what is probably the most flexible way of any nonhuman signal, since both when they are produced and the ways in which they are aimed at particular addressees indicate sensitivity to relatively complex social dynamics (Armstrong, 2019; Seyfarth and Cheney, 2015). Still, these signals can function only in stereotyped social contexts, and their content must pertain to current social dynamics. Nonhuman primates do not communicate about past, future, or non-actual events, and the topics of their exchanges are highly circumscribed.

By contrast, humans communicate about an astounding range of topics, many of which have nothing to do with the immediate situation in which we find ourselves. No inherent limitation in the human communicative capacity prevents a stranger from approaching you in a public restroom to tell you about what the latter-day music of Jimi Hendrix would have been like if he hadn't died young, for example. In general, although there are some stereotyped situations and topics of human communication, there are no principled limits on when and where we communicate, or on which things—actual or possible, past, present or future—we communicate about. And although natural language makes available specialized grammatical resources for displacing the topic of conversation from the here, the now, and the actual (von Fintel and Heim, 2011), this phenomenon is by no means limited to linguistic communication. There is a whole genre of children's books that tells fictional stories using only pictures, for example, and we are capable of doing the same with pantomime in the right circumstances (Sibierska, 2017).

## **9 Our ability to design signals for our addressees**

Humans interpret signals in flexible ways, and also make flexible decisions about when, where, and what to communicate. But we also make highly flexible decisions after we've decided what to communicate, when it comes time to design a signal with which to communicate it to a particular addressee. We customize our signals in ways that no other creatures do.

Suppose that you want to tell me something about a particular person, and you need to choose a way of referring to them in order to do so. You have many referring expressions to choose from, and making a good selection will require accounting for the situation in which we're talking, including your information about the information that is available to me. One possibility is that the person you want to tell me about is me, and you know this, in which case you should utter 'you'. Another possi-

bility is that it's someone else nearby whom you judge to be salient from my perspective, in which case you should say 'he', 'she', or 'they'. Otherwise, you might want to use their name, but you should do this only if you take me to know their name. If all else fails, you could formulate a description of the person. But this involves you in a whole new set of decisions, as there are an infinite number of true descriptions of any person, only some of which will provide me with a useful way of identifying them. The best models of how speakers choose referring expressions suggests that they do make use of information about their addressee in approximately the ways I've just described, often quickly and unconsciously.<sup>10</sup>

This is just one of the many ways that we design both our linguistic and non-linguistic signals with particular addressees in mind. It would not be unusual for a Swiss person to speak their local dialect of Swiss German at home, French at school, high German at work, and English at the train station. If someone asks me to explain my research, I will use rather different vocabulary depending on what I take to be their level of interest and expertise. This can be quite fine-grained: my choice of technical vocabulary could depend on something as specific as which semantics textbook I take my addressee to have studied. The same goes for much nonlinguistic communication: A savvy communicator will know that they should shake their head from side to side in order to refuse when in Rome, but should toss their chin up in the air in order to refuse when in Naples (Morris, 1977). In a diagram for a contractor whom they know to be red-green colorblind, an architect will avoid using those colors to communicate important distinctions. If, while driving somewhere with my wife, I want to signal that she should stop what she is doing and enjoy some particularly beautiful scenery, I will play the third movement of Philip Glass's fifth String Quartet. I know she will understand me only because of idiosyncratic details of our shared history—this is a signal designed just for her.

Our ability to customize signals to particular addressees makes possible some of our most advanced communicative tricks. We sometimes design a signal to have different effects on different members of our audience, knowing that their different background information will lead them to understand the signal in different ways. Political scientists and philosophers have recently studied the way that politicians use this ability to communicate different messages to different segments of their audience (Khoo, 2017; Mendelberg, 2001; Saul, 2018). An example used by Saul

---

<sup>10</sup>See, for example, Clark and Marshall (1981); Clark and Murphy (1982); Clark et al. (1983); Dale and Reiter (1995); Degen et al. (2019); Hawkins et al. (2019); Horton and Brennan (2016); Keysar (2007). I will discuss this evidence in detail in Chapter Three.

(2018) involves George Bush's use of the phrase 'wonder-working power' during his 2003 State of the Union speech. The phrase carries no special significance for most Americans, but it is commonly used by evangelical preachers to refer to the powers of Jesus, and so served as a calculated way for Bush to emphasize his evangelical identity to those who shared it. This sort of audience-specific double meaning is not limited to those who can afford to hire political consultants. Scott (1990) makes an extended case that slaves, serfs, untouchables, and other oppressed peoples often use songs, folk tales, jokes, theater, and religion as media in which to covertly communicate about and formulate methods of resisting their oppression while in plain sight of their oppressors. For example, American slaves developed a genre of Christianity that emphasized the emancipation stories of Moses and Joshua, which served as a form of coded sedition, but one that was shielded from slaveowners behind the pious veil of their own religion (Scott, 1990, 158).

There is little evidence of audience-specific signal design by nonhuman organisms. Once again, the best analogy is with our fellow primates, and is quite limited. Primate vocalizations have sometimes been described as involuntary and audience-unspecific—as “individualistic expressions of emotions, not recipient-directed acts” (Tomasello, 2008, 19). This isn't entirely correct: as we have already seen, many primates issue alarm calls in ways that are sensitive to whether they have an audience at all, and whether the audience is already aware of the threat. Unlike primate alarm calls, which are addressed to any other group members who are present, great-ape gestures and affiliative and aggressive vocalizations are addressed to specific individuals, and apes check to make sure that they have the addressee's visual attention, sometimes intentionally moving into their line of sight before gesturing (Tomasello and Call, 2019). However, there are serious limits on how primates manage to design their signals for their addressees. For example, they notably do not intentionally use attention-getting gestures as a way of drawing attention to other gestures (Liebal et al., 2004). And primates do not in any way customize the sound patterns of vocalizations for different audiences (Seyfarth and Cheney, 2018). Primate signal customization is therefore limited to a decision about whether or not to send a signal at all on the basis of whether they have an (attentive) audience.

## 10 Our ability to engage in organized communicative exchanges

Individual instances of human communication usually occur as parts of extended, organized communicative exchanges—paradigmatically, conversations. Consider this example:

- (3) (i) ANN: I met a very tall man today whom I think you'd like.
- (ii) BEA: Why do you think I would like him?
- (iii) ANN: He shares your passion for lepidopterology.
- (iv) BEA: You should invite this towering butterfly lover to your party.
- (v) ANN: Maybe he'll bring a net!
- (vi) BEA: It would would make him hard to miss.

This brief slice of dialogue illustrates several of the mechanisms that we use to knit together our linguistic acts into larger conversations. We ask questions in order to solicit responses of specific kinds from our addressees, as in (ii)–(iii) (Carlson, 1982; Roberts, 2012). We use indefinite noun phrases, such as ‘a very tall man’ in (i), to introduce new individuals as topics of conversation, and definite noun phrases, such as the pronouns found throughout (ii)–(vi), to refer back to and say more about them (Heim, 1983; Karttunen, 1976). Similarly, we use some modals, such as ‘maybe’ in (v), to introduce new possibilities as topics, and other modals, such as ‘would’ in (vi), to indicate that we are further elaborating on possibilities that have already been introduced (Roberts, 1989; Stone, 1997). Utterances are often designed to be interpreted in ways that presuppose information introduced earlier in the conversation (Beaver and Geurts, 2014). In order to identify the referent of ‘this towering butterfly lover’ in (iv), for example, Ann needs to synthesize information gleaned from (i) and (iii) about his height and interest in butterflies, respectively.

These and other grammatical phenomena comprise the connective tissue of conversations, making them more than the sum of their parts. Although they make any one contribution intelligible only in the context of others, this allows individual utterances to convey much more than they could in isolation. And it allows us to gradually build toward complex communicative goals that can only be accomplished using cumulative inquiries or slow-burning narratives.

Although the exemplary communicative exchanges are conversations, in which much of the work is done by special-purpose grammatical devices, there are also

cruder nonlinguistic examples. Imagine two drivers fighting over a parking spot from within their sealed-off cars. They don't speak sign language, and so they are forced to rely on a series of gestures, facial expressions, and honks. Still, we can imagine them progressing through a conversation-like exchange, with the parking spot serving as a continuous topic: driver A points at the spot and then at his eyes to signal that it was his first, driver B shakes his head to signal disagreement and then taps his watch to signal that he is in a hurry, A crosses his arm to signal recalcitrance and shrugs his shoulders as if to inquire how long B is willing to hold out, and so on. We can imagine progress being made: perhaps B will hold up a \$20 bill to offer a bribe, and A will nod to accept it. Here we have an extended exchange in which each element is understandable because of the context provided by the others, and in which the state of A's and B's interaction slowly evolves with the performance of a variety of communicative acts, including assertions, questions, commands, and offers. Like many of the properties of human communication on this list, our ability to assemble our communicative acts into extended exchanges is heightened by language's grammatical affordances, but the underlying phenomenon is not essentially linguistic.

There are some nonhuman forms of communication that resemble our communicative exchanges in some respects. Many animals respond to each other's signals, but this normally takes the form of repeating the same signal type—or, in the case of songbirds, signals with the same functions. Likewise, some primates continue producing alarm calls repeatedly until all of their group members have responded, but the responses always take the form of the same alarm call repeated back, serving to indicate only that the respondents have been successfully alerted (Wich and de Vries, 2006). Perhaps the most conversation-like phenomena found in nonhuman animals involves the affiliative and aggressive signals of great apes. In some species, the apes in a given group continually keep track of who has recently been addressing aggressive or friendly signals to whom, and they use this information gleaned from past signals in order to interpret later signals. Armstrong (2019) argues that this phenomenon supports an analysis that in some ways resembles the standard theoretical treatments of presupposition and anaphora in human language. However, there are severe limitations on how these ape conversations can be said to work. Signals have a small number of functions, consisting only of signaling aggression or friendliness toward other group members, and information gathered from earlier signals is needed to interpret later signals only insofar as it is needed to identify their addressees. This is somewhat more sophisticated than it may sound,



since it relies on apes' representations of complex patterns of social relations, but it is a far cry from humans' communicative exchanges. Neither apes nor any other nonhuman organisms can do anything analogous to asking a followup question (or any question at all), saying a series of different things about a specific third party, or elaborating on a possibility raised earlier.

## 11 The varieties of human communication

I have pointed out ten ways that our capacity for communication is markedly more powerful and flexible than the analogous capacities of other organisms.<sup>11</sup> Of course, many instances of human communication lack some of these features. Most of our nonlinguistic signals lack anything like compositional structure. Our pheromonal signals, like those of other organisms, are limited to a fixed and unlearned range of chemical compounds and communicative functions, cannot be used indirectly, cannot be interpreted in light of background information, can be sent only in a physiologically proscribed range of circumstances, and cannot be customized to particular addressees. In addition to the rich and flexible communication that makes us uniquely human, we also communicate in ways that resemble what other primates do. At least some facial expressions have the same communicative functions across all humans, are prompted by specific emotional states in ways that are under limited voluntary control, and normally cannot be used indirectly or in ways that require addressees to draw on much background information. The Duchenne smile is one example.

Our capacity for communication is thus a palimpsest, with more sophisticated mechanisms layered on top of simpler and evolutionarily older ones. This should not be surprising: evolution needn't remove old traits when it adds new traits with similar functions. We therefore shouldn't expect there to be any one way that humans communicate.

One might object to the foregoing that I have been cherry-picking—treating our greatest communicative accomplishments as though they are run-of-the-mill human behaviors. If this objection were on target, then what I have done in this chapter is akin to characterizing our capacity for locomotion by saying that humans can run marathons in about two hours, or describing our strength by pointing out that humans can deadlift over 500kg. While strictly-speaking true, these statements are misleading, or at least interesting only as descriptions of rare virtuosity rather

---

<sup>11</sup>I reiterate that I do not claim that this list is exhaustive.

than normal human behavior. But my intention has not been to catalogue only our most bizarre and impressive communicative feats. Rather, I take myself to have described ten entirely normal and everyday features of human communicators—things that most of us repeatedly manifest on an average day.

## 12 Looking Forward

What explains the fact that human communication is unique in these ways? This book presupposes that an answer to this question will be a theory of the psychological capacities that add up to our capacity to communicate, and of the psychological mechanisms that underlie these capacities. My aim is to develop such a theory—one about which I am as optimistic as it could be reasonable to be, given the present state of our scientific understanding of human beings. It will take the whole book to present and defend the theory. In the next chapter, I will give an overview of it and show how it explains why our capacity to communicate, but not those of other organisms, possess the characteristics that I have outlined here.

## References

- Abney, S. P. (1987). *The English Noun Phrase in its Sentential Aspect*. PhD thesis, MIT, Cambridge, MA.
- Armstrong, J. (2019). Provincialism in pragmatics. *Philosophical Perspectives*, Early View.
- Arnold, K. and Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or idiomatic expressions? *Brain and Language*, 120(3):303–309.
- Austin, J. L. (1962). *How to do Things with Words*. The Clarendon Press, Oxford.
- Azhar, H. (2016). How we really use the peach. In *Emojiopedia*. URL = <<https://blog.emojiopedia.org/how-we-really-use-the-peach/>>.
- Bach, K. and Harnish, R. M. (1979). *Linguistic Communication and Speech Acts*. MIT Press, Cambridge, Mass.
- Barwise, J. and Perry, J. (1983). *Situations and Attitudes*. MIT Press.
- Beaver, D. and Geurts, B. (2014). Presupposition. In Zalta, E. N., editor, *Stanford Encyclopedia of Philosophy (Winter 2014 Edition)*. URL = <<https://plato.stanford.edu/archives/win2014/entries/presupposition/>>.
- Berwick, R. C., Okanoya, K., Beckers, G. J. L., and Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences*, 15(3):113–121.
- Bloom, P. (2000). *How Children Learn the Meanings of Words*. MIT Press, Cambridge, MA.
- Buchanan, K. L. and Catchpole, C. K. (1997). Female choice in the sedge warbler, *Acrocephalus schoenobaenus*: multiple cues from song and territory quality. *Proceedings of the Royal Society B*, 264:521–526.
- Byun, K.-S., de Vos, C., Bradford, A., Zeshan, U., and Levinson, S. C. (2018). First encounters: Repair sequences in cross-signing. *Topics in Cognitive Science*, 10:314–334.
- Camp, E. (2006). Metaphor and that certain ‘je ne sais quoi’. *Philosophical Studies*, 129(1):1–25.
- Carey, S. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, 15:17–29.
- Carlson, L. (1982). *Dialogue Games: An Approach to Discourse Analysis*. Number 17 in Synthese Language Library. D. Reidel, Dordrecht, The Netherlands.
- Chomsky, N. (2002). *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Blackwell, Oxford.

- Catchpole, C. K. (1973). The functions of advertising song in the sedge warbler (*acrocephalus schoenobaenus*) and the reed warbler (*a. scirpaceus*). *Behaviour*, 46(3/4):300–320.
- Catchpole, C. K. (1976). Temporal and sequential organization of song in the sedge warbler (*acrocephalus schoenobaenus*). *Behaviour*, 59(3/4):226–246.
- Catchpole, C. K. (1980). Sexual selection and the evolution of complex songs among european warblers of the genus *acrocephalus*. *Behaviour*, 74(1/2):149–166.
- Catchpole, C. K., Dittami, J., and Leisler, B. (1984). Differential responses to male song repertoires in female songbirds implanted with oestradiol. *Nature*, 312:563–564.
- Catchpole, C. K. and Slater, P. J. B. (2008). *Bird Song: Biological Themes and Variations*. Cambridge University Press, Cambridge, second edition.
- Cheney, D. L. and Seyfarth, R. M. (1988). Assessment of meaning and the detection of unreliable signals by vervet monkeys. *Animal Behavior*, 36:477–486.
- Cheney, D. L. and Seyfarth, R. M. (1990). *How Monkeys See the World*. University of Chicago Press, Chicago.
- Cheney, D. L. and Seyfarth, R. M. (1999). Recognition of other individuals' social relationships by female baboons. *Animal Behavior*, 58:67–75.
- Cheney, D. L. and Seyfarth, R. M. (2007). *Baboon Metaphysics: The Evolution of a Social Mind*. University of Chicago Press, Chicago and London.
- Cheney, D. L. and Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, 115(9):1974–1979.
- Clark, E. (2009). *First Language Acquisition*. Cambridge University Press, Cambridge, second edition.
- Clark, H. H. and Marshall, C. R. (1981). Definite reference and mutual knowledge. In Joshi, A. K., Webber, B. L., and Sag, I. A., editors, *Elements of Discourse Understanding*, pages 10–63. Cambridge University Press, Cambridge.
- Clark, H. H. and Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, 9:287–299.
- Clark, H. H., Schreuder, R., and Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, 22:245–258.
- Dale, R. and Reiter, E. (1995). Computational interpretations of the gricean maxims in the generation of referring expressions. *Cognitive Science*, 18:233–263.
- Degen, J., Hawkins, R. D., Graf, C., Kreiss, E., and Goodman, N. D. (2019). When redundancy is useful: A bayesian approach to 'overinformative' referring expressions. *arXiv*, arXiv:1903.08237:1–82.
- Dyer, F. C. and Dickenson, J. A. (1994). Development of sun compensation by honeybees: How partially experience bees estimate the sun's course. *Proceedings of the National Academy of Sciences*, 91:4471–4474.
- Engh, A. L., Hoffmeier, R. R., Cheney, D. L., and Seyfarth, R. M. (2006). Who, me? can baboons infer the target of vocalizations? *Animal Behavior*, 71:381–387.
- von Fintel, K. (1994). *Restrictions on Quantifier Domains*. PhD thesis, University of Massachusetts, Amherst, Amherst, MA.
- von Fintel, K. and Heim, I. (2011). *Intensional Semantics*. Unpublished Lecture Notes, online at <http://web.mit.edu/fintel/fintel-heim-intensional.pdf>, spring 2011 edition.
- Fitch, W. T. (2016). Why formal semantics and primate communication make strange bedfellows. *Theoretical Linguistics*, 42(1–2):97–109.
- von Frisch, K. (1927). *Aus dem Leben der Bienen*. Springer-Verlag, 1927.
- Gardner, R. A. and Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, 165(3894):664–672.
- Greenfield, P. M. and Savage-Rumbaugh, E. S. (1990). Grammatical combination in pan paniscus: Processes of learning and invention in the evolution and development of language. In Parker, S. T., editor, *'Language' and Intelligence in Monkeys and Apes*, chapter 20, pages 540–578. Cambridge University Press, Cambridge.
- Habib, A. (2018). Promises. In Zalta, E. N., editor, *Stanford Encyclopedia of Philosophy (Spring 2018 Edition)*. <<https://plato.stanford.edu/archives/spr2014/entries/promises/>>.
- Harms, W. F. (2004a). *Information and Meaning in Evolutionary Processes*. Cambridge University Press, Cambridge.
- Harms, W. F. (2004b). Primitive content, translation, and the emergence of meaning in animal communication. In Oller, D. K. and Griebel, U., editors, *Evolution of Communication Systems: A Comparative Approach*, The Vienna Series in Theoretical Biology, chapter 3, pages 31–48. MIT Press, Cambridge, MA.
- Harris, D. W., Fogal, D., and Moss, M. (2018). Speech acts: The contemporary theoretical landscape. In Fogal, D., Harris, D. W., and Moss, M., editors, *New*

- Work on Speech Acts*. Oxford University Press, Oxford.
- Hawkins, R. D., Gweon, H., and Goodman, N. D. (2019). The division of labor in communication: Speakers help listeners account for asymmetries in visual perspective. *arXiv*, arXiv:1807.09000:1–43.
- Heil, M. and Karban, R. (2010). Explaining evolution of plant communication by airborne signals. *Trends in Ecology and Evolution*, 25(3):137–144.
- Heim, I. (1983). File change semantics and the familiarity theory of definiteness. In Bäuerle, R., Schwarze, C., and von Stechow, A., editors, *Meaning, Use and Interpretation of Language*, pages 164–189. de Gruyter, Berlin.
- Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell, Oxford.
- Horton, W. S. and Brennan, S. E. (2016). The role of metarepresentation in the production and resolution of referring expressions. *Frontiers in Psychology*, 7:1–12.
- Hurford, J. R. (2012). *The Origins of Grammar*, volume 2 of *Language in the Light of Evolution*. Oxford University Press, Oxford.
- Janik, V. M., Sayigh, L. S., and Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences*, 103(21):8293–8297.
- Jensen, M. N. (2000). Whales’ cultural revolution. *Science*.
- Kako, E. (1999). Elements of syntax in the systems of three language-trained animals. *Animal Learning and Behaviour*, 27(1):1–14.
- Kaminski, J., Call, J., and Fischer, J. (2004). Word learning in a domestic dog: Evidence for “fast mapping”. *Science*, 304(5677):1682–1683.
- Karttunen, L. (1976). Discourse referents. In McCawley, J. D., editor, *Syntax and Semantics 7: Notes from the Linguistic Underground*, pages 363–385. Academic Press.
- Kaufmann, M. (2012). *Interpreting Imperatives*. Springer, Dordrecht.
- Kennedy, C. (2007). Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy*, 30:1–45.
- Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, 4(1):71–84.
- Kho, J. (2017). Code words in political discourse. *Philosophical Topics*, 45(2):33–64.
- King, S. L. and Janik, V. M. (2013). Bottlenose dolphins can use learned vocal labels to address each other. *Proceedings of the National Academy of Sciences*, 110(32):13216–13221.
- Knight, C. (2001). Does cultural evolution need matriline. *Behavioral and Brain Sciences*, 24:339–340.
- Kratzer, A. (1981). The notional category of modality. In Eikmeyer, H. J. and Rieser, H., editors, *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. de Gruyter, Berlin.
- Kroodtsma, D. E. and Parker, L. D. (1977). Vocal virtuosity in the brown thrasher. *The Auk*, 94(4):783–785.
- Levinson, S. C. (2006). On the human “interaction engine”. In Levinson, S. C. and Enfield, N. J., editors, *Roots of human sociality: Culture, cognition and interaction*. Berg, Oxford.
- Liebal, K., Call, J., and Tomasello, M. (2004). Use of gesture sequences in chimpanzees. *American Journal of Primatology*, 64:377–398.
- Lindauer, M. (1961). *Communication among Social Bees*. Harvard University Press, Cambridge, MA.
- Marco, R. D. and Menzel, R. (2005). Encoding spatial information in the waggle dance. *Journal of Experimental Biology*, 208:3885–3894.
- Marler, P. (1998). Animal communication and human language. In Jablonski, N. G. and Aiello, L. C., editors, *The Origin and Diversification of Language*, pages 1–19. California Academy of Sciences, San Francisco, CA.
- Melville, H. (1852). *Pierre; or, The Ambiguities*. Harper and Brothers, New York.
- Mendelberg, T. (2001). *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton University Press, Princeton, NJ.
- Miller, M. B. and Bassler, B. L. (2001). Quorum sensing in bacteria. *Annual Review of Microbiology*, 55:165–199.
- Millikan, R. G. (1995). Pushmi-pullyu representations. *Philosophical Perspectives*, 9:185–200.
- Morris, D. (1977). *Manwatching*. Jonathan Cape, London.
- Neale, S. (1990). *Descriptions*. MIT Press.
- Neale, S. (2016). Silent reference. In Ostertag, G., editor, *Meanings and Other Things: Essays in Honor of Stephen Schiffer*, pages 229–342. Oxford University Press, Oxford.
- Noad, M. J., Cato, D. H., Bryden, M. M., Jenner, M.-N., and Jenner, K. C. S. (2000). Cultural revolution in whale songs. *Nature*, 408:537.
- Ouattara, K., Lemasson, A., and Zuberbühler, K. (2009). Campbell’s monkeys use affixation to alter call meaning. *PLoS ONE*, 4(11):e7808.

- Papineau, D. (2001). The evolution of means-end reasoning. *The Royal Institute of Philosophy Supplement*, 49:145–178.
- Patterson, F. G. (1978). The gestures of a gorilla: Language acquisition by another pongid. *Brain and Language*, 5:72–97.
- Patterson, F. G. P. and Matevia, M. L. (2001). Twenty-seven years of project koko and michael. In Gaaldikas, B. M. F., Briggs, N. E., Sheeran, L. K., Shapiro, G. L., and Goodall, J., editors, *All Apes Great and Small: African Apes*, pages 165–176. Springer, New York.
- Pepperberg, I. M. (1999). *The Alex Studies*. Harvard University Press, Cambridge, MA.
- Pilley, J. W. (2013). Elements of syntax in the systems of three language-trained animals. *Learning and Motivation*, 44(4):229–240.
- Pilley, J. W. and Reid, A. K. (2011). Border collie comprehends object names as verbal referents. *Behavioural Processes*, 86(2):184–195.
- Pinker, S. (1994). *The Language Instinct*. William Morrow and Company, New York.
- Portner, P. (2004). The semantics of imperatives within a theory of clause-types. In Watanabe, K. and Young, R., editors, *Proceedings of SALT 14*. CLC Publications.
- Rasmussen, L. E. L., Lee, T. D., Roelofs, W. L., Zhang, A., and Jr, G. D. D. (1996). Insect pheromone in elephants. *Nature*, 379:684.
- Reznikova, Z. (2007). Dialog with black box: using information theory to study animal language behaviour. *Acta Ethologica*, 10(1):1–12.
- Reznikova, Z. and Ryabko, B. (1986). Analysis of the language of ants by information-theoretical methods. *Problems of Information Transmission*, 22(3):245–249.
- Rivas, E. (2005). Recent use of signs by chimpanzees (*Pan Troglodytes*) in interactions with humans. *Journal of Comparative Psychology*, 119(4):404–417.
- Roberts, C. (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12(6):683–721.
- Roberts, C. (2012). Information structure in discourse: Toward an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5:1–69.
- Saul, J. (2018). Dogwhistles, political manipulation, and the philosophy of language. In Fogal, D., Harris, D. W., and Moss, M., editors, *New Work on Speech Acts*, pages 360–383. Oxford University Press, Oxford.
- Savage-Rumbaugh, S. and Lewin, R. (1994). *Kanzi: The Ape at the Brink of the Human Mind*. John Wiley and Sons, New York.
- Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., Stephan, C., Ryder, R., and Zuberbühler, K. (2014). Monkey semantics: Two ‘dialects’ of Campbell’s monkey alarm calls. *Linguistics and Philosophy*, 37(6):439–501.
- Schlenker, P., Chemla, E., Arnold, K., and Zuberbühler, K. (2016a). *Pyow-hack* revisited: Two analyses of putty-nosed monkey alarm calls. *Lingua*, 171:1–23.
- Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gauthier, J.-P., Kuhn, J., Veselinovi, D., Arnold, K., Cäsar, C., Keenan, S., Lemasson, A., Ouattara, K., Ryder, R., and Zuberbühler, K. (2016b). Formal monkey linguistics. *Theoretical Linguistics*, 42(1–2):1–90.
- Scott, J. C. (1990). *Domination and the Arts of Resistance*. Yale University Press, New Haven, CT.
- Scott-Phillips, T. C., Kirby, S., and Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition*, 113:226–233.
- Seidenberg, M. S. and Petitto, L. A. (1979). Signing behavior in apes: A critical review. *Cognition*, 7(2):177–215.
- Seyfarth, R. and Cheney, D. (2018). Pragmatic flexibility in primate vocal production. *Current Opinion in Behavioral Sciences*, 21:56–61.
- Seyfarth, R. M. and Cheney, D. L. (2015). The evolution of concepts about agents: Or, what do animals recognize when they recognize an individual? In Margolis, E. and Laurence, S., editors, *The Conceptual Mind: New Directions in the Study of Concepts*, pages 57–75. MIT Press, Cambridge, MA.
- Sibierska, M. (2017). Storytelling without telling: The non-linguistic nature of narratives from evolutionary and narratological perspectives. *Language & Communication*, 54:47–55.
- Silk, J. B. (2001). Grunts, girneys, and good intentions: The origins of strategic commitment in nonhuman primates. In *Evolution and the Capacity for Commitment*, pages 138–158. Russell Sage Foundation, New York.
- Silk, J. B., Kaldor, E., and Boyd, R. (2001). Cheap talk when interests conflict. *Animal Behavior*, 59:423–432.
- Slater, P. J. B. (1986). The cultural transmission of bird song. *Tree*, 1(4):94–97.
- Sperber, D. and Wilson, D. (1995). *Relevance: Communication and Cognition*. Blackwell, Oxford.

- Stone, M. (1997). The anaphoric parallel between modality and tense. *IRCS Technical Reports Series*, pages 1–44.
- Terrace, H. S., Petitto, L. A., Sanders, R. J., and Bever, T. G. (1979). Can an ape create a sentence? *Science*, 206(4421):891–902.
- Tomasello, M. (2008). *Origins of Human Communication*. MIT Press.
- Tomasello, M. and Call, J. (2019). Thirty years of great ape gestures. *Animal Cognition*, 22:461–469.
- Wallman, J. (1992). *Aping Language*. Cambridge University Press, Cambridge.
- Waring, G. H. (1970). Sound communications of black-tailed, white-tailed, and gunnison's prairie dogs. *The American Midland Naturalist*, 83(1):167–185.
- Whitehead, H. and Rendell, L. (2015). *The Cultural Lives of Whales and Dolphins*. University of Chicago Press, Chicago and London.
- Wich, S. A. and de Vries, H. (2006). Male monkeys remember which group members have given alarm calls. *Proceedings of the Royal Society B*, 273:735–740.
- Wyatt, T. D. (2014). *Pheromones and Animal Behavior: Chemical Signals and Signature Mixtures*. Cambridge University Press, Cambridge, second edition.
- Yang, C. (2013). Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences*, 110(16):6324–6327.
- Zollman, K. J. S. (2011). Separating directives and assertions using simple signaling games. *The Journal of Philosophy*, 108(3):158–169.
- Zuberbühler, K. (2001). Predator-specific alarm calls in campbell's monkeys, *Cercopithecus campbelli*. *Behavioral Ecology and Sociobiology*, 50(5):414–422.