JOURNAL OF THE ROYAL SOCIETY Interface

Why is combinatorial communication rare in the natural world, and why is language an exception to this trend?

Thomas C. Scott-Phillips and Richard A. Blythe

J. R. Soc. Interface 2013 **10**, 20130520, published 18 September 2013

JOURNAL OF THE ROYAL SOCIETY Interface

rsif.royalsocietypublishing.org

Research

Cite this article: Scott-Phillips TC, Blythe RA. 2013 Why is combinatorial communication rare in the natural world, and why is language an exception to this trend? J R Soc Interface 10: 20130520.
http://dx.doi.org/10.1098/rsif.2013.0520

Received: 13 June 2013
Accepted: 14 August 2013

Subject Areas:
computational biology

Keywords:
animal signals, combinatorial communication, communication, evolution, language, ostension

Author for correspondence:
Thomas C. Scott-Phillips
e-mail: t.c.scott-phillips@durham.ac.uk

†These authors contributed equally to this study.

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsif.2013.0520 or via http://rsif.royalsocietypublishing.org.

Royal Society Publishing
Informing the science of the future

Why is combinatorial communication rare in the natural world, and why is language an exception to this trend?

Thomas C. Scott-Phillips[1,†] and Richard A. Blythe[2,†]

[1]Evolutionary Anthropology Research Group, Department of Anthropology, Durham University, Dawson Building, South Road, Durham DH1 3LE, UK
[2]SUPA, School of Physics and Astronomy, University of Edinburgh, James Clerk Maxwell Building, Mayfield Road, Edinburgh EH9 3JZ, UK

In a combinatorial communication system, some signals consist of the combinations of other signals. Such systems are more efficient than equivalent, non-combinatorial systems, yet despite this they are rare in nature. Why? Previous explanations have focused on the adaptive limits of combinatorial communication, or on its purported cognitive difficulties, but neither of these explains the full distribution of combinatorial communication in the natural world. Here, we present a nonlinear dynamical model of the emergence of combinatorial communication that, unlike previous models, considers how initially non-communicative behaviour evolves to take on a communicative function. We derive three basic principles about the emergence of combinatorial communication. We hence show that the interdependence of signals and responses places significant constraints on the historical pathways by which combinatorial signals might emerge, to the extent that anything other than the most simple form of combinatorial communication is extremely unlikely. We also argue that these constraints can be bypassed if individuals have the socio-cognitive capacity to engage in ostensive communication. Humans, but probably no other species, have this ability. This may explain why language, which is massively combinatorial, is such an extreme exception to nature's general trend for non-combinatorial communication.

1. Introduction

In a combinatorial communication system, some signals consist of the combinations of other existing signals. The most basic version involves two individual signals that are combined to refer to something that is not simply the amalgamation of whatever the two individual signals refer to, but something different. In other words, combinatorial communication includes at least one *composite* signal, whereas in non-combinatorial communication, all the signals are *holistic* (figure 1). For example, putty-nosed monkeys are reported to have two distinct alarm calls, one for each of two predators: leopards (a 'pyow' sound) and eagles (a 'hack' sound) [1,2]. When one or the other of these calls is produced on its own, the monkeys take appropriate evasive action: climbing up and into the trees for leopards; climbing down and into the bushes for eagles. However, when the two calls are produced together ('pyow–hack'), the effect is not the simple combination of these, i.e. the monkeys do not behave as if avoiding both types of predator. Instead, the call seems to presage the movement of the group to a new location (perhaps, for example, because there is a shortage of fresh food at the present location).

Combinatorial communication has one obvious adaptive advantage over equivalent non-combinatorial systems: fewer elements are required to express the same number of possible messages, and so it allows for more efficient communication than a system in which each signal has a distinct form [3]. Despite this potential advantage, it is rare in nature, and where it does exist it is, with one salient exception, simple and limited [4,5]. Many systems (e.g. honeybee

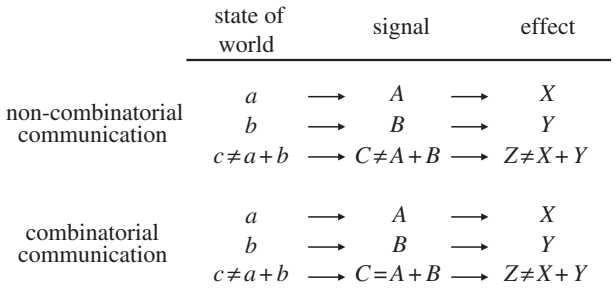

Figure 1. Combinatorial communication. In a combinatorial communication system, two (or more) *holistic* signals (*A* and *B* in this figure) are combined to form a third, *composite* signal ($A + B$), which has a different effect (Z) to the sum of the two individual signals ($X + Y$). This figure illustrates the simplest combinatorial communication system possible. Applied to the putty-nosed monkey system, the symbols in this figure are: *a*, presence of eagles; *b*, presence of leopards; *c*, absence of food; *A*, 'pyow'; *B*, 'hack' call; $C = A + B$ 'pyow–hack'; *X*, climb down; *Y*, climb up; $Z \neq X + Y$, move to a new location. Combinatorial communication is rare in nature: many systems have a signal $C = A + B$ with an effect $Z = X + Y$; very few have a signal $C = A + B$ with an effect $Z \neq X + Y$.

dance) have a signal which has an effect that is equal to the sum of its component parts, but very few have a signal whose effect is different to this sum (in the terms of figure 1, $Z = X + Y$ is common, but $Z \neq X + Y$ is not). The salient exception is human language, which is massively combinatorial. Indeed, there are multiple different types of combination that can contribute to meaning. Otherwise, the only well-attested examples are the putty-nosed monkey case described above.

Why, then, is combinatorial communication so rare, and why is language such an extreme exception? One possible explanation might be that combinatorial communication is cognitively challenging in some way [6,7]. However, why this should be the case is unclear: there is no particular reason to think that signal combinations should be difficult to process. Other previous explanations have focused on the limits of the adaptive benefits of combinatorial communication. One analysis shows that the benefits associated with combinatorial communication are realized only when the total number of signals in the system exceeds a threshold level [3,8]. However, this prediction is not supported by the data: several systems have many more signals than this threshold level, but still do not combine them (non-human primate gestural communication, for example [9]); and the putty-nosed monkey system described above has fewer signals than this threshold level, but is combinatorial nevertheless. Finally, one other analysis concludes that combinatorial systems are more susceptible to dishonesty than non-combinatorial systems [10]. However, this does not explain why language should be such a clear exception.

In this paper, we develop a new explanation of the distribution of combinatorial communication in nature. The historical pathways by which traits evolve are an important source of constraint on biological form [11], yet most models of the evolution of communication, including all those focused on combinatorial communication, ask only how one type of communicative system can evolve from another, and not how an initially non-communicative behaviour can take on a communicative function. Specifically, previous models do not include strategies for signaller and receiver prior to the emergence of a signal, despite the fact that such prior strategies are likely to have considerable impact on the form of the eventual communication system [12,13]. We present a deterministic, nonlinear dynamical model that includes such strategies. We use this to derive a number of principles regarding the origins of composite signals, and hence show that the emergence of composite signals is subject to significant historical constraints, to the extent that even the most simple forms of combinatorial communication are likely to be uncommon, and anything more complex vanishingly so. We then argue that these constraints can be bypassed if a species has the social cognitive abilities to communicate ostensively, i.e. in a way that involves the expression and recognition of communicative and informative intentions. Humans, but probably no other species, have this ability.

2. The emergence of communication

Before we present the model, it is necessary to briefly describe the two classic ways in which communication systems can emerge. In general, new signals emerge through either ritualization or sensory manipulation (also called sensory exploitation) [13–15]. In ritualization, previously existing cues are exapted for use as signals (a cue is a behaviour that is informative for other organisms, but was not selected to be so [14,16]). For example, the use of urine to mark territorial boundaries probably first began when animals urinated simply through fear, when they were at the edge of familiar territory. This acted as a cue to other animals, who make use of that information. This in turn provides a selection pressure on the focal organism to urinate when and if it wants/needs to inform others about the range of its territory. In sensory manipulation, previously existing behaviour is exapted for use as a response. For example, male scorpionflies capture large prey and then offer them to females who feed on them during copulation [17]. The offering of prey by the male probably initially evolved, because the female had a pre-existing mechanism that prioritized the opportunity to feed on large prey, and so the presentation of food gave the male an opportunity to mate. There was then later positive selection on the female to accept the prey in exchange for copulation [14].

These processes, ritualization and sensory manipulation, constrain the form that new signals can take. Our previous model showed that if a particular behaviour does not exist for reasons independent of communication prior to either of these processes occurring, then it cannot become a signal or response, regardless of its adaptive value for signaller or receiver [13]. Prospective signals and responses must already provide fitness benefits to either receiver or signaller, respectively, if they are to actually evolve into signals/responses. In other words, there must be some sort of trigger, external to the (proto-)communicative interaction (i.e. either a cue or a coercive behaviour), to cause a signal to actually emerge. With this background in place, we now develop a formal model of the emergence of specifically combinatorial communication systems.

3. Basic set-up of model, and classification of communication systems

The environment can be in any one of a number of different states, $E \in \mathscr{E}$. Given this state, one agent (the actor) performs

an action, $A(E) \in \mathscr{A}$. Another agent (the reactor) then performs a reaction, $R(A) \in \mathscr{R}$. Pay-offs are determined by the different combinations of states and reactions. This much is the same as the standard game-theoretic approach to modelling communication. However, we differ from the standard approach in our specification of the sets $\mathscr{E}$, $\mathscr{A}$ and $\mathscr{R}$. In particular, in our model, these sets include default settings (E_0, A_0 and R_0), which correspond to the agents doing nothing, and which are orthogonal to all other members of the set. Typically, these default settings are not included in game-theoretic models of communication. However, they are critical to understanding how communication can emerge from a state of non-communication [13].

With the exception of the defaults E_0 and A_0, environments and actions can be combined with other environments and actions. We denote these composites as $E_i \circ E_j$ and $A_i \circ A_j$, respectively. These composites are, like their component parts, members of $\mathscr{E}$ and $\mathscr{A}$ respectively, and composition is commutative (i.e. order does not matter, so $E_i \circ E_j = E_j \circ E_i$ and $A_i \circ A_j = A_j \circ A_i$).

$A(E)$ and $R(A)$ are deterministic functions of E and A, respectively. Together, they comprise an agent's strategy. If there are non-composite environments E_i and E_j in which the actions A_i and A_j are performed, then the composite action $A_i \circ A_j$ is performed in the composite environment $E_i \circ E_j$. Two agents are members of the same species if they share the same functions A and R for all possible environments and actions, respectively (i.e. two individuals i and j are members of the same species $A_i(E) = A_j(E)\ \forall\ E$ and $R_i(A) = R_j(A)\ \forall\ A$).

A signal is any non-default action that yields a non-default reaction, and any such reaction is called a response. A communication system is a set of more than one signal–response pair. Within a communication system, each given pair of actions A_i and A_j (where both $\neq A_0$) can be classified in one of three ways:

— *Non-composite*: a pair is non-composite if the composite of the two actions is produced only in composite environments, and it, in turn, yields the default reaction, i.e. there is no $E_k \neq E_i \circ E_j$ such that $A(E_k) = A_i \circ A_j$, and $R(A_i \circ A_j) = R_0$.
— *Pseudo-composite*: a pair is pseudo-composite if the composite of the two actions is produced only in composite environments, and it, in turn, yields a non-default reaction, i.e. there is no $E_k \neq E_i \circ E_j$ such that $A(E_k) = A_i \circ A_j$, while at the same time $R(A_i \circ A_j) \neq R_0$.
— *Fully-composite*: a pair is fully-composite if the composite of the two actions is produced in at least one non-composite environment, and it in turn yields a non-default reaction, i.e. $\exists\ E_k \neq E_i \circ E_j$ such that $A(E_k) = A_i \circ A_j$, and $R(A_i \circ A_j) \neq R_0$. A combinatorial communication system is a system that includes at least one pair of fully-composite actions.

(Logically, there is a fourth possible class, where the composite of two actions is produced in at least one non-composite environment, and it, in turn, yields the default reaction. However, such a pair is unstable, because the default reaction produces only a zero pay-off for the actor, by definition, and this is outweighed by the maintenance and production costs of the composite pair. We therefore ignore this possibility in the subsequent analysis.)

Communication is vulnerable to instability caused by dishonesty. How communication systems remain stable in the face of this problem is an important and much studied question for the evolution of communication [14,18]. Here, however, we are concerned with a different question, namely assuming that communication is evolutionarily stable, what are the different evolutionary pathways by which (combinatorial) communication can emerge? We thus wish to avoid issues of stability, which might complicate our analysis, and so we assume that at least one of the different mechanisms that can stabilize communication is in place. In particular, we find that ascribing a direct benefit for successful communication to both signaller and receiver [as in e.g. 19,20], or imposing kin discrimination (i.e. agents can observe the actions of their conspecifics only), leads to the same set of dynamical equations (see electronic supplementary material). We expect other possible mechanisms to lead to the same or similar principles for the emergence of combinatorial communication as those we set out below.

4. Dynamics

We now derive the dynamics for the model. The basic principles of the model are that: the frequency of a particular environment is given by $f(E)$; if an agent performs a reaction R in the environment E, then there is a pay-off $s(R|E)$; the actor must also receive a pay-off, either directly or indirectly, for communication to be stable (see above); and there is a cost associated with each agent's strategy, i.e. the cost of having the capacity to behave in a non-default way in the first place (this is χ for actions and η for reactions; see below). This is distinct from the cost associated with each individual behaviour, which we include as part of the pay-off $s(R|E)$. These two types of cost can be thought of as *maintenance costs* and *energy costs*, respectively, and the inclusion of both is an important difference between our model and previous models of the emergence of communication [19,20].

We first write down the dynamical equations for the frequencies of the various strategies in the population, and we define $\psi(A,E)$ as the fraction of agents who perform action A in the environment E, and $\phi(R,A)$ as the fraction of agents who perform reaction R in response to action A. As we show in the electronic supplementary materials, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi(A,E) = \psi(A,E)\left[u(A,E) - \sum_{A'} u(A',E)\psi(A',E)\right],$$

where

$$u(A,E) = f(E)\sum_{R} s(R|E)\phi(R,A) + \sum_{E' \neq E,E_0} f(E \circ E') \sum_{R} s(R|E \circ E') \times \sum_{A'} \psi(A',E)\phi(R,A \circ A') + \chi\delta(A,A_0) \qquad (4.1)$$

(δ is the Kronecker delta symbol, which equals 1 if the two arguments are identical, and 0 otherwise.) The equation has the same structure as the replicator equation in the standard evolutionary game theory [21], in which $u(A,E)$ is the fitness of the rule $E \rightarrow A$. There are three contributions to this fitness:

— The first term gives the fitness of the rule $E \rightarrow A$ in the environment E, given the current distribution of the possible rules $R(A)$.

— The second term gives the net fitness of the rule $E \rightarrow A$ in all the composite environments that include E, given the current distribution of the possible rules $R(A)$. This term is unique to composite signals. It has the consequence that the spontaneous emergence of composite signalling strategies (either actions or reactions) can be favourable.
— The third term accounts for the cost, χ, of having a mechanism for producing non-default actions.

Similarly, for the frequency $\phi(R,A)$ of the rule $A \rightarrow R$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(R,A) - \phi(R,A)\left[v(R,A) - \sum_{R'} v(R',A)\psi(R',A)\right],$$

where

$$v(R,A) = \sum_{E} f(E)s(R|E)\psi(A,E) + \sum_{\langle E_1,E_2\rangle} f(E_1 \circ E_2)s(R|E_1 \circ E_2) \times \sum_{(A_1,A_2)} \psi(A_1,E_1)\psi(A_2,E_2)\delta(A_1 \circ A_2, A) + \eta\delta(R,R_0). \tag{4.2}$$

As above, there are three terms to this equation, which correspond to: the fitness of the rule $A \rightarrow R$ in non-composite environments; the average fitness of a composite action $A = A_1 \circ A_2$ performed in composite environments; and the cost, η, of having a mechanism for producing non-default reactions. Also as above, we derive this equation step-by-step in the electronic supplementary material.

These equations can be applied to any specific set of components, i.e. states of the environment; possible actions and reactions; and parameters. In the electronic supplementary material, we define these components for a specific model, in order to test the general predictions we derive below. Its results are entirely consistent with the principles we set out below.

5. Three basic principles for the emergence of combinatorial communication systems

We now derive three basic principles that govern how a combinatorial communication system might emerge. We are interested, primarily, in the case where two actions, A_i and A_j, are composed to become the action that is used to signal an elementary environmental state E_k that is unrelated to E_i and E_j. We will not discuss higher-order composite actions (i.e. those where one or both components are themselves composite actions), but we have no reason to think that the general principles we derive here should be any different in that case.

5.1. Principle 1: only homogeneous populations are evolutionarily stable

For any given environmental state, the fitness, $u(A,E)$, of each action rule is independent of the frequencies, $\psi(A',E)$, of its competitors. The action rule with the highest fitness will then grow at the expense of all its competitors until it is the only one remaining. If two or more rules have the same fitness as each other, then drift will likewise eliminate all but one of them. These observations also apply to reaction rules. Hence, only homogeneous populations, in which all agents have the same set of rules, are evolutionarily stable. We thus assume homogeneous populations in what follows.

5.2. Principle 2: new non-composite signals cannot emerge without an external trigger

Given some pre-existing signalling system, a new non-composite signal is the action, A, in the set of rules $E \rightarrow A \rightarrow R$ ($A \neq A_0$, $R \neq R_0$), where (i) there is currently no environment in which the new action A is performed; (ii) A is not a composite of any two existing actions in the system and (iii) the existing reaction to A, produced by itself or in combination with any other action A', is the default reaction. In a homogeneous population (see principle 1), condition (i) implies that the first term in equation (4.2) vanishes for both $v(R,A)$ and $v(R_0,A)$ because $\psi(A,E) = 0$ for all E; condition (ii) implies that the second term in equation (4.2) also vanishes for both $v(R,A)$ and $v(R_0,A)$, because $\delta(A_i \circ A_j, A) = 0$ for all A_i, A_j; and condition (iii) implies that the first two terms of equation (4.1) are the same for both $u(A,E)$ and $u(A_0,E)$ because $\phi(R',A) = \phi(R',\ A \circ A') = \delta(R',R_0)$ for all A'. Consequently, for any new non-composite signalling behaviour, we always have that

$$u(A,E) - u(A_0,E) = -\chi \text{ in any environment,}$$
$$v(R,A) - v(R_0,A) = -\eta.$$

In other words, adding either the action or reaction component of this signalling behaviour carries a cost for every individual in the population. Hence, a completely new signalling behaviour that involves an action that is not currently part of the signalling system, cannot emerge without some external trigger as described in §2. This result is an extension of our previous result, that a signal cannot be added to an existing state of non-communication without an external trigger (cue or coercion) [13]. The observation here is that this issue also applies to the addition of a any non-composite signal to an existing communication system. This point will be important when we discuss human linguistic communication, in §7, below.

5.3. Principle 3: new composite signals can emerge without an external trigger

Consider a pair of existing signals, A_i and A_j. If there is an environment where the combination of these two signals would provide a cue (useful information) for other organisms, then their co-production can lead to the evolution of a composite signal. More formally, if the environments E_i and E_j trigger the actions A_i and A_j, respectively, then the composite environment $E_i \circ E_j$ will trigger the composite action $A_i \circ A_j$. Repeating the analysis from principle (2), we find, again, that adding a new action will be costly for all individuals in the population, i.e. that $u(A_i \circ A_j,\ E_k) - u(A_0,\ E_i \circ E_j) = -\chi$ for all E, as before. As with principle (2), this prevents the emergence of a composite signal by sensory manipulation. However, unlike principle (2), the emergence of a composite signal by ritualization is possible. That is, there may be a reaction that is not yet an existing reaction to either A_i or A_j, and which would be beneficial for the receiver to perform in the composite environment, i.e. it is possible for $v(R,A) - v(R_0,A) > 0$ for some R. If this condition is satisfied, then, assuming that the relevant environments occur sufficiently often for evolution to occur, we should expect the corresponding reaction to evolve. We will then have arrived at a pseudo-composite signal:

$(E_i \circ E_j) \rightarrow (A_i \circ A_j) \rightarrow R \neq R_0$. The key observation here is that the co-production of two existing signals can itself be the trigger required for a new signal to emerge. This possibility is absent in the case of non-composite signals. Note that from here, it is then possible for a new rule $E_k \neq (E_i \circ E_j) \rightarrow (A_i \circ A_j)$ to emerge by sensory manipulation, giving us a fully-composite signal. A concrete demonstration of this possibility is given in the electronic supplementary material, where we apply our model to the specific case of putty-nosed monkey alarm calls. Note, however, that the emergence of a full-composite signal is not guaranteed (it is possible, for example, that reactions to higher-order compositions could prevent this).

6. Why combinatorial communication should be rare

What do these principles imply for the emergence of a combinatorial communication system? One immediate observation is that they explain how fully-composite signals can emerge even in a simple world of just two existing signals. This is noteworthy because it is contrary to a previous analysis, which argued that composite signals should only emerge within systems of multiple (more than five) signals [3]. However, the present empirical data suggest, consistent with our analysis, that composite signals can exist even in very simple systems (see, e.g. the putty-nosed monkey system described in §1).

Another observation might be that the three principles derived above seem to imply that composite signals should be far more common than non-composite signals. After all, composite signals can emerge without an external trigger (principle 3), but non-composite signals cannot (principle 2). However, this reading fails to take account of the conditions attached to each of these possibilities. There are two in particular that we wish to highlight.

The first is the relative frequency by which the various triggers of the emergence of composite and non-composite signals occur. In particular, although the emergence of a composite signal does not require a trigger external to the system itself, it does require one from within the system. This condition is sufficiently stringent that the emergence of a composite signal is in fact less likely to occur than the external triggers that are required for the emergence of non-composite signals. Here is why. The internal trigger required for the emergence of a new composite signal is a very specific one: that, given the co-production of two existing signals, A_i and A_j, then there is a reaction R that (i) if it were performed in the composite environmental state $E_i \circ E_j$, it would be beneficial to the receiver; and (ii) that this reaction is not yet an existing reaction to either A_i or A_j (see principle 3). In other, more information-centric terms, what is required is that the co-production of two existing signals must be informative about some aspect of the world, beyond what can be deduced from the meanings of the individual signals themselves—and there is no particular reason why this should be the case. By contrast, a new non-composite signal can emerge from *any* behaviour that an individual might perform (see §2). In other words, there is one specific way that any new signal might be composite, but a vast number of ways, limited only by the number of behaviours the organism can actually perform, that any new signal might be non-composite. Consequently, composite signals should be rare. This is not to say that they cannot emerge, only that their emergence is dependent on unlikely prior circumstances. Hence, they should be rare, at least in comparison with non-composite signals.

The second condition attached to the emergence of fully-composite signals is that it is dependent on the instability of other possible systems. Consider a basic system of $E_i \rightarrow A_i \rightarrow R_i$ and $E_j \rightarrow A_j \rightarrow R_j$ (i.e. just the first two signals in figure 1). Principle 3 states that it is then possible for a fully-composite signal $E_k \rightarrow (A_i \circ A_j) \rightarrow R_k$ to emerge, to form the system described in figure 1, without an external trigger. However, it turns out that this is only true if the alternative system, of the basic system plus a holistic signal $E_k \rightarrow A_k \rightarrow R_k$ is unstable. Here is why. In order for the process described in principle 3 to occur, the basic system must be unstable to the addition of $(A_i \circ A_j) \rightarrow R_k$. A necessary condition for this to be the case is that in this system, $v(R_k, A_i \circ A_j) > v(R_0, A_i \circ A_j)$. Using the equation for $v(R,A)$ in §4, we find that for this instability to be present, we require

$$f(E_i \circ E_j)[s(R_k|E_i \circ E_j) - s(R_0|E_i \circ E_j)] - \eta > 0. \tag{6.1}$$

However, using the same equation for $v(R,A)$, but now applied to the alternative system (i.e. the one that includes $E_k \rightarrow A_k \rightarrow R_k$ rather than $E_k \rightarrow (A_i \circ A_j) \rightarrow R_k$), we find that for this alternative system to be stable we require

$$f(E_i \circ E_j)[s(R_k|E_i \circ E_j) - s(R_0|E_i \circ E_j)] - \eta < 0. \tag{6.2}$$

Equations (6.1) and (6.2) clearly contradict each other. This shows that the conditions required for the process described in principle 3 to occur include that the alternative system is evolutionarily unstable (note that this is true whether or not the fully-compositional system described in figure 1 is evolutionarily stable). As such, this is an additional criterion on the emergence of composite signals, and hence on the emergence of combinatorial communication.

In sum, there are at least two conditions that can work to restrict the emergence of composite signals. The first is that the triggers required for the emergence of composite signals are less likely to occur than are the triggers for non-composite signals. The second is that the process of emergence without an external trigger depends on the instability of any alternative, holistic system. Both these conditions are the consequence of the interdependence of signals and responses, and they help to explain why combinatorial communication is rare in nature.

7. Human linguistic communication

There is, of course, one extreme exception to the norm of non-combinatorial communication: human linguistic communication. Here, meaningless sounds (phonemes) are combined into meaningful units (morphemes), which are, in turn, combined into utterances, whose meaning is a function not only of the morphemes involved, but also the order in which they are combined (a feature called *duality of patterning*: [22,23]). This combinatorial richness gives language its expressive power [5,24]. How can we explain why language is such a clear exception to the general trend for non-combinatorial systems? In this section, we use the conclusions from our model to pinpoint and articulate an important difference between human and animal communication. We hence argue that human linguistic communication is simply not subject to the various historical contingencies described

in the previous sections—and consequently, combinatorial communication is free to emerge wherever it may be useful.

Human communication depends, at bottom, on mechanisms of metapsychology: that is, the ability to reason about others' reasons, intentions, beliefs and so on. Communication of this sort is called *ostensive communication* [25]. Linguistic communication is an instance of ostensive communication that has been made expressively powerful by the development of rich suite of communicative conventions that allow it to be used far more precisely and expressively than it otherwise would [26]. A signaller can, for example, ostensively point to any of the objects in this room, but with language she can refer to any object in the world. She also can make a request of others by, for example ostensively pushing unchopped vegetables, and a knife, in their direction, but with language she can make requests about things remote in time and space. Other examples are not hard to imagine. By contrast, most, and perhaps all, animal communication depends on mechanisms of association: causal relationships between stimuli and responses (but see below). Communication of this sort is called *coded communication* (see references [25–27] for discussion of the difference between coded and ostensive communication).

Our model in this paper has been a model of coded communication: we have studied how states of the world become *associated* with certain actions, and how these actions, in turn, become *associated* with certain reactions. Indeed, all models of animal communication that study the emergence of such associations are code models. However, such models do not capture an important fact about ostensive communication: that meaning is not deduced or calculated, even probabilistically, on the back of associations (be they between signal and meaning, or perhaps between signals, context and meaning), but rather it is inferred, based on the receiver's beliefs about the signaller's intentions [25]. This inference is, unlike the associations that make coded communication possible, made possible by metapsychology [27].

One consequence of this difference is that human ostensive communication, including linguistic communication, is inherently prone to ambiguity. This is generally seen as a defective quality, because it can, on occasion, lead to misunderstanding and other failures of communication. However, it also allows communication to be used in flexible, creative and open-ended ways—and these ways include the combination of already existing signals. One consequence of this is that the spaces of possible signal forms and signal meanings become continuous, rather than become discrete. This development is possible only because signallers have the metapsychological abilities to create the right sort of signal to express their intended meaning, whatever it might be, and because receivers have similar abilities to infer those intended meanings.

Here is an example. Homesigners are deaf children born to hearing parents. Lacking the input of a conventional sign language, they must create new communication systems themselves, and this includes the combination of existing signals [28]. Here is one very simple case [28]. The child, Karen, is already familiar with pointing, and also with a 'twist' gesture that means 'open'. She then uses these two behaviours together: she points to a jar of soap bubbles and then, without pausing, produces an iconic 'twist' action with her hands. In doing so, she indicates to the adult that she would like her to open the jar. At first blush, this seems unremarkable, but that is only because, as fluent users of ostensive communication, we are fully accustomed to such acts of creation as an everyday occurrence. The point here is not that we can combine things together. It is rather that, because she has the required metapsychological abilities, it is possible for Karen to provide just the right sort of evidence, given her intended meaning and her intended audience. This is ostensive communication. It just happens that in this case the right sort of evidence happens to involve the combination of two existing signals.

Note that in a different context, the meaning of Karen's behaviour could be very different indeed. Suppose, for example, that the adult had just tried to open the jar by twisting it, but had failed, and that this had amused Karen. Now Karen could use the same combined signal to make a humorous reference to this past event. This flexibility is possible only because both Karen and the adult have the metapsychological abilities required. On the ostensive side, Karen produced the signals in such a way that it was apparent that they are in fact one signal, comprised of two parts; this is why she does not pause between the two. On the inferential side, the adult must assess what Karen's intended meaning was, given her knowledge of the context, and of the meanings of the two component parts.

Here is how this example relates to our model. Karen has an existing set of actions that she produces in particular environments, and these receive particular responses from the adult. Specifically

E_1 = Karen wants to refer to an out-of-reach object;
A_1 = pointing;
R_1 = attention is focused in the direction of the point;
E_2 = Karen wants the adult to open something;
A_2 = 'twist' gesture; and
R_2 = the adult opens the object of mutual attention.

Karen finds herself in a new environment: E_3 = Karen wants the adult to open an out-of-reach object. Note that this environment is *not* the sum of the other two: $E_3 \neq E_1 \circ E_2$. Instead, the composite environment $E_1 \circ E_2$ is the co-occurrence of (i) an object that is out-of-reach object; and (ii) an object that Karen wishes to open. There is nothing in this that specifies that these two objects should in fact be the same object: that aspect of the scenario is additional, and as such is specific to E_3. Our model shows that, without an external trigger to set the evolutionary process in motion, it is not possible, in a coded communication system, for a new, non-composite signal such as this to emerge (principle 2, above). Yet here, not only does such a signal emerge, it does so immediately, and smoothly: there is no interruption of the normal flow of communication. Neither is this an instance of the emergence of communication by ritualization, in which a cue evolves into a signal (see principle 3, above)—because Karen's behaviour is not a cue. It is a signal from the moment of its production, and that is the point. As such, this is a clear exception to the general constraints described previously. In sum, the existence of ostensive communication makes it possible for a species to overcome the constraints, described above, that otherwise make the emergence of combinatorial communication unlikely.

There is, then, an important sense in which Karen's twist signal contrasts with superficially similar signals in a coded communication system. Coded 'combinatorial' signals are in a sense not really combinatorial at all. After all, there is no 'combining' going on. There is really just a third holistic signal, which happens to be comprised of the same pieces

as other existing holistic signals. Indeed, the most recent experimental results suggest that the putty-nosed monkeys interpret the 'combinatorial' pyow–hack calls in exactly this idiomatic way, rather than as the product of two component parts of meaning [29]. By contrast, the ostensive creation of new composite signals is clearly combinatorial: the meaning of the new, composite signal is in part (but only in part) a function of the meanings of the component pieces.

It is presently unclear whether any other species uses ostensive communication. The precise psychological mechanisms necessary are cognitively complex, and so it is quite possible that it is uniquely human [30–32]. Certainly, this would be consistent with the argument we have developed in this paper, and there is presently no convincing evidence that any other species communicates ostensively [30,32]. However, this remains, at least for now, an open empirical question. (Note that ostensive communication is not the same thing as intentional communication, which some other species certainly do use.)

8. Conclusion

Previous models of the emergence of combinatorial communication were focused on the following question: under what circumstances are composite signals advantageous, in comparison with holistic signals? Our model in this paper addresses a different question: by what processes can composite signals emerge? To do this, we explicitly modelled the possibility that no communication might take place: this is why our model includes the default states E_0, A_0 and R_0, which are absent from other models. Our results show that combinatorial communication is rare in nature, because the interdependence of signals and responses constrains the ways by which communication systems emerge, with the effect that novel signals will tend to be holistic rather than tend to be composite. However, this constraint can be bypassed if the communication system in question is ostensive—and this type of communication is likely unique to humans. Unlike other proposals (see Introduction), this explanation is consistent with all the empirical facts: it explains both why combinatorial communication is generally rare in the natural world, and why there is a single, extreme exception to this trend.

Acknowledgements. We thank Robert Barton for comments on a previous draft.

Funding statement. T.C.S.P. acknowledges financial support from the Leverhulme Trust and the ESRC, and R.A.B. from Research Councils UK.

References

1. Arnold K, Zuberbühler K. 2006 Semantic combinations in primate calls. *Nature* **441**, 303. (doi:10.1038/441303a)
2. Arnold K, Zuberbühler K. 2008 Meaningful call combinations in a non-human primate. *Curr. Biol.* **18**, R202–R203. (doi:10.1016/j.cub.2008.01.040)
3. Nowak MA, Plotkin JB, Jansen VA. 2000 The evolution of syntactic communication. *Nature* **404**, 495–498. (doi:10.1038/35006635)
4. Fitch WT. 2010 *The evolution of language*. Cambridge, UK: Cambridge University Press.
5. Hurford JR. 2011 *The origins of grammar*. Oxford, UK: Oxford University Press.
6. Zuberbühler K. 2002 A syntactic rule in forest monkey communication. *Anim. Behav.* **63**, 293–299. (doi:10.1006/anbe.2001.1914)
7. Ouattara K, Lemasson A, Zuberbühler K. 2009 Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proc. Natl Acad. Sci. USA* **106**, 22 026–22 031. (doi:10.1073/pnas.0908118106)
8. Nowak MA, Krakauer DC. 1999 The evolution of language. *Proc. Natl Acad. Sci. USA* **96**, 8028–8033. (doi:10.1073/pnas.96.14.8028)
9. Hobaiter C, Byrne RW. 2011 The gestural repertoire of the wild chimpanzee. *Anim. Cogn.* **14**, 745–767. (doi:10.1007/s10071-011-0409-2)
10. Lachmann M, Bergstrom CT. 2004 The disadvantage of combinatorial communication. *Proc. R. Soc. Lond. B* **271**, 2337–2343. (doi:10.1098/rspb.2004.2844)
11. Ridley M. 2009 *Evolution*. Oxford, UK: Wiley.
12. Bradbury JW, Vehrencamp SL. 2000 Economic models of animal communication. *Anim. Behav.* **59**, 259–268. (doi:10.1006/anbe.1999.1330)
13. Scott-Phillips TC, Blythe RA, Gardner A, West SA. 2012 How do communication systems emerge? *Proc. R. Soc. B* **279**, 1943–1949. (doi:10.1098/rspb.2011.2181)
14. Maynard Smith J, Harper DGC. 2003 *Animal signals*. Oxford, UK: Oxford University Press.
15. Bradbury JW, Vehrencamp SL. 1998 *Principles of animal communication*. Sunderland, MA: Sinauer Associates.
16. Scott-Phillips TC. 2008 Defining biological communication. *J. Evol. Biol.* **21**, 387–395. (doi:10.1111/j.1420-9101.2007.01497.x)
17. Thornhill R. 1976 Sexual selection and nuptial feeding in *Bittacus apicalis* (Insecta: Mecoptera). *Am. Nat.* **110**, 529–548. (doi:10.1086/283089)
18. Searcy WA, Nowicki S. 2007 *The evolution of animal communication*. Princeton, NJ: Princeton University Press.
19. Donaldson MC, Lachmann M, Bergstrom CT. 2007 The evolution of functionally referential meaning in a structured world. *J. Theor. Biol.* **246**, 225–233. (doi:10.1016/j.jtbi.2006.12.031)
20. Jäger G. 2008 Evolutionary stability conditions for signaling games with costly signals. *J. Theor. Biol.* **253**, 131–141. (doi:10.1016/j.jtbi.2008.02.039)
21. Nowak MA. 2006 *Evolutionary dynamics: exploring the equations of life*. Harvard, NJ: Belknap Press.
22. Hockett CF. 1969 The origin of speech. *Sci. Am.* **203**, 88–111. (doi:10.1038/scientificamerican0960-88)
23. Ladd DR. 2012 What is duality of patterning, anyway? *Lang. Cogn.* **4**, 261–273.
24. Chomsky N. 1965 *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
25. Sperber D, Wilson D. 1995 *Relevance: communication and cognition*, 2nd edn. Oxford, UK: Blackwell.
26. Origgi G, Sperber D. 2010 A pragmatic perspective on the evolution of language. In *The evolution of human language: biolinguistic perspectives* (eds RK Larson, V Déprez, H Yamakido), pp. 124–132. Cambridge, UK: Cambridge University Press.
27. Sperber D, Wilson D. 2002 Pragmatics, modularity and mind-reading. *Mind Lang.* **17**, 3–23. (doi:10.1111/1468-0017.00186)
28. Goldin-Meadow S. 2005 *The resilience of language*. Abingdon, UK: Psychology Press.
29. Arnold K, Zuberbühler K. 2012 Call combinations in monkeys: compositional or idiomatic expressions? *Brain Lang.* **120**, 303–309. (doi:10.1016/j.bandl.2011.10.001)
30. Tomasello M. 2008 *Origins of human communication*. Cambridge, MA: MIT Press.
31. Fitch WT, Huber L, Bugnyar T. 2010 Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* **65**, 795–814. (doi:10.1016/j.neuron.2010.03.011)
32. Scott-Phillips TC. Submitted. Non-human primate communication, pragmatics, and the origins of language.