'Speaking Our Minds' Book Club

The International Cognition and Culture Institute

June-July 2015

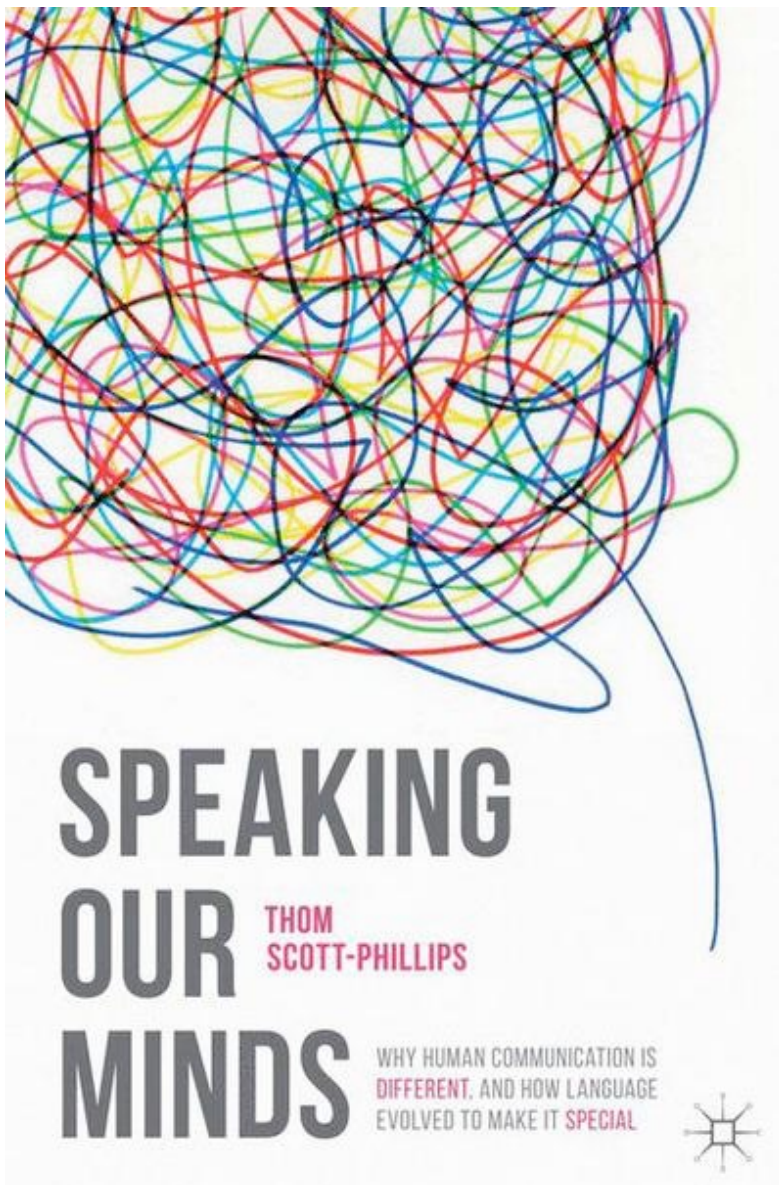
From mid-June to mid-July 2015, the ICCI website has hosted a Book Club devoted to Thom Scott-Phillips' book, Speaking Our Minds. The book has received positive reviews in various media, including the Times Literary Supplement (read Richard Moore's review here), and has been called "an amazing job" (Stuart West),"the most important and best book ever written on the evolution of language" (Dan Sperber).

TABLE OF CONTENTS

1

A PRÉCIS OF 'SPEAKING OUR MINDS'

By Thom Scott-Phillips



"Speaking our Minds: Why Human Communication is Different and How Language Evolved to Make it Special" Palgrave Macmillan, 2014 (212 pages).

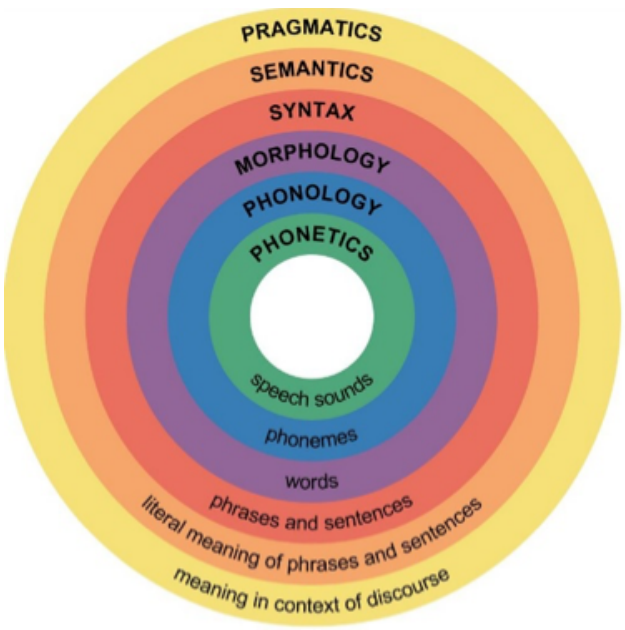
Communication and language have always been key topics for research at the interface of cognition and culture. Rightly so, given the central role that linguistic communication plays in human social and cultural life. In fact, communication and language are doubly important, since they occupy both sides of the cognition and culture coin. On the one side is a cognitive ability, to engage in linguistic communication in the first place. On the other side are cultural objects, namely languages themselves, which are collections of communicative conventions shared within a population. This double importance is reflected in the ambiguity of the phrase "language evolution", which is used both to describe the study of how humans evolved to communicate in the ways that they do, and the study of how languages themselves evolve, culturally, to take the various forms that they do.

In 1866 the *Société de Linguistique de Paris* declared that it would no longer consider correspondence on the topic of language origins ("La Société n'admet aucune communication concernant, soit l'origine du langage – soit la création d'une langue universelle"). This "ban" is often said to have curtailed language evolution research, only for it to be reawakened in the 1990s, but this cli-

ched history has little truth to it. A great deal of language evolution research occurred between 1866 and 1990. Darwin himself speculated on the origins of language just a few years after the Parisian edict, and several 20th-century research agendas directly address language origins. The clearest and most well-known example is the many attempts to teach human language to non-human apes. Contrary to common assertion, 1990 was not year zero for language evolution research.

It is however fair to say that it was during the 1990s that something like an academic field of language evolution began to emerge. Researchers from different disciplines began to share ideas, the first conferences focused on the topic took place, and some pioneering individuals began to consider it their main research interest. This growth has continued since. We should not exaggerate the extent of this, but we can reasonably say that there is now a field of language evolution, it is here to stay, there is some outstanding science done in its name (some less than outstanding science too), and also that the various disciplines concerned with language and communication are increasingly open to evolutionary perspectives. This is most clear within linguistics itself. For many years mainstream linguistics was at best uninterested in evolutionary approaches, and at worst actively hostile to them. This is no longer the case, as a browse of the major journals of the discipline will quickly show. It is no longer unusual to find work that is informed by an evolutionary perspective in one way or another.

This influence of language evolution on mainstream linguistics has been mirrored by a corresponding influence in the other direction, although the impact of this is more subtle, and less easy to observe. For most of its history as an academic discipline, linguistics' principal intellectual concern has been with the form and structure of languages. The classic divisions between different levels of language structure are nicely illustrated in this figure:



Major levels of linguistic structure
Source: Wikipedia

All but the outermost circle of this figure describe items – sounds, phonemes, words, phrases, and literal meaning – that are either discrete in character, or can be treated as such. The sixth, outermost circle is where the linguistic rubber hits the communicative road, and this interface with the outside world brings with it a raft of issues that complicate linguistic analysis. Depending on one's specific questions, it can be productive to abstract away from these issues, but the unintended consequence of doing so is that, despite good intentions and well-meaning lip-service, pragmatics is kept at the margins of linguistics. The fact that it is concerned with the boundary between language and the outside world is, ironically, a major reason why pragmatics is kept on the periphery of the discipline itself.

This marginalisation has affected the direction of language evolution research, and continues to do so. If you doubt this, turn to the index of the *Oxford Handbook of Language Evolution* (Tallerman & Gibson, 2011), a book billed (accurately) as "a wide-ranging summation of work in all the disciplines involved [in language evolution]". There you will find 213 pages listed under 'syntax' and related terms, 100 pages listed under 'phonetics', 'phonology' and related terms, but just 8 pages under 'pragmatics'. Or, alternatively, take a look at the lists of plenary speakers from the ten Evolang conferences that have taken place to date. You will find a relative dearth of pragmaticists. Two especially conspicuous omissions, particularly in comparison with the relevance of some who have delivered Evolang plenaries, are Stephen Levinson and Dan Sperber, both high-profile pragmaticists who have written extensively about the origins and evolution of human communication and language.

This neglect of pragmatics is, I believe, a profound mistake. For some fields (theoretical syntax, say) and for some questions, abstraction away from the messy realities of human social interaction, and towards the more abstract, idealised world of discrete linguistic items, can be a reasonable and profitable agenda. However, I do not believe that language evolution is such a field, at least not in general, and this is true whether one is concerned with its cognitive or its cultural dimensions. How languages are used in communication is, instead, critical to any evolutionary explanation of why humans are able to communicate in this way, and of why languages themselves take the forms that they do. If I am right about this, then rather than being peripheral, pragmatics should be a key concern of language evolution – and indeed of cognition and culture studies in general. One way to read *Speaking Our Minds* is as a demonstration of just how much we can learn about language evolution by taking pragmatics seriously. Put simply, it is my belief that with a pragmatic perspective, it is possible to develop cogent, compelling answers to all the major questions asked about language evolution (and I think this is true even if one does not agree with the specifics of my own answers).

Let me expand a little, and in doing so provide a little background for my own claims. In *Relevance: Communication and Cognition* (Sperber & Wilson, 1986/1995), not to mention many publications since, Dan Sperber and Deirdre Wilson argued that linguistic communication exists on a continuum with other, non-linguistic forms of communication, canonical examples of which include points, shrugs, and other non-verbal gestures. The overall category here is that of *ostensive communication*:

communication that involves the expression and recognition of intentions. These intentions are, specifically, communicative intentions (which we can roughly gloss as an intention to make apparent to the audience *that* one is trying to communicate) and informative intentions (which we can roughly gloss as an intention to make apparent to the audience *what* one is trying to communicate). Ostensive communication can be used for a great many communicative ends (there is much you can do just with non-linguistic grunts and gestures), but its expressivity is hugely increased by the addition of words, grammar, and the other communicative conventions that collectively comprise a language. I can make a request of others by ostensively pushing unchopped vegetables in their direction, but with specific conventions I can make requests about things remote in time and space. Linguistic communication is, then, a special case of ostensive communication, namely one in which expressivity is hugely increased by the existence of shared communicative conventions.

If all this is correct, then already a pragmatic perspective has earned its keep, since it makes clear what the two most general questions for language evolution should be. First, how and why did humans evolve ostensive communication (and do any other species communicate ostensively?). Second, how do collections of communicative conventions develop, and how and why do they transition towards the forms that they do? The short versions of my answers to these questions are that ostensive communication is uniquely human, it evolved as a by-product of enhanced social intelligence, conventions emerge through interaction and use, and they gravitate towards particular forms through a process of cultural attraction, in order to most closely fit the goals and dispositions of language users. These answers collectively motivate my subtitle: *Why Human Communication Is Different, & How Language Evolved To Make It Special.*

Chapter 1: Two Approaches to Communication

In the first chapter of *Speaking Our Minds*, I distinguish ostensive communication from code model communication. Code model communication is achieved through pairs of association: one between a state of the world and a signal, and the other between the signal and a response. Following arguments developed by Sperber and Wilson in Relevance Theory, I defend the claim that ostensive communication cannot be reduced to pairs of associations, and hence that the difference between ostensive communication and code model communication is not one of degree, but one of kind. More precisely, the difference between these two types of communication is at bottom a difference in the sort of mechanism that makes communication possible in the first place.

Ostensive communication is made possible by mechanisms of social cognition, whereas code model communication is made possible by mechanisms of association. (Whether or not any species other than humans communicates ostensively, or whether all animal communication is code model communication, is an empirical question that I defer until chapter four.)

Where does this distinction between ostensive and code model communication leave language and linguistic communication? Clearly associations are involved in some way: there is a linguistic code. However, these associations do not function to make linguistic communication possible in the first place (and hence linguistic communication does not operate according to the code model). As I discussed above, linguistic associations (words, grammar, etc.) instead function to make ostensive communication more precise and more expressive than it otherwise would be. I adopt the labels *natural codes* and *conventional codes* to distinguish between, respectively, codes that function to make communication possible in the first place, and codes that function to make a different type of communication more expressive. Languages are not natural codes, but instead (sets of) conventional codes.

Chapter 2: The Emergence of Communication Systems

How do different types of communication systems emerge? I argue in chapter two that the process is quite different for the two different types of communication distinguished in chapter one.

Most work on the origins of communication systems focuses, quite reasonably, on code model communication. Here, there is a chicken-and-egg problem: which comes first, the signal or the response? After all, one without the other is pointless, and natural selection does not act with foresight. This problem may be especially serious for the emergence of combinatorial codes, in which two existing signals are combined to achieve an effect that is different to the sum of the effects of the two component parts. The obvious solution to the chicken-and-egg problem is that one of the two behaviours first evolves for reasons independent of communication, and this then provides the right selective environment for the other. The processes of emergence described in the animal communication literature, both in theory and in the empirical data, match this prediction.

Ostensive communication is quite different. The ability to express and recognise intentions means that any behaviour can, in principle, be used communicatively, so long as it is produced in an ostensive way. This means that there is no chicken-and-egg problem to contend with, and new signals can be created as and when required. Some of my own experimental work shows this process in action, but it can also be observed in natural data, such as in the emergence of homesign (novel systems created by deaf children of hearing parents).

Chapter 3: Cognition and Communication

I said above that ostensive communication is made possible by mechanisms of social cognition (by definition it involves the expression and recognition of intentions). Chapter three discusses these mechanisms in more detail. I begin by coining the term *pragmatic competence*: the ability to use ostensive communication in a competent way. As signallers, this is the ability to produce the right sort of behaviour to express one's intended meaning; as listeners, it is the ability to make the right sort of in-

ferences about that behaviour. The question is: what cognitive mechanisms make pragmatic competence possible? In answering this question, I take as my starting point Relevance Theory, which I believe provides a *bona fide* scientific paradigm for pragmatics. I argue in particular that recursive mindreading – the ability not only to infer what others believe (think, desire, etc.), but also to infer what others believe about what further others believe – is a necessary component of any computational description of what is involved in ostensive communication.

There is a tension between, on the one hand, the theoretical cogency of these arguments, and on the other, the intuition that recursive mindreading is a cognitively demanding activity, certainly too demanding to be involved in everyday communicative interaction. To date, most researchers who have grappled with this problem have chosen to accept the intuition, or some version of it, and argue against the theory. That is to say, they have tried to develop arguments that ostensive communication in fact requires (supposedly) less demanding cognitive resources. I adopt the other approach: I accept the theory, and argue against the intuition that recursive mindreading is cognitively demanding. In particular, I discuss how these arguments can be squared with the data on the mindreading abilities of children, who demonstrate pragmatic competence from a young age.

Chapter 4: The Evolution of Ostensive Communication

Does any other species communicate ostensively? Do human children? There is a sizeable literature on intentional communication in other species, but intentional communication is not, I believe, the same thing as ostensive communication. Intentional communication involves signals being used in a goal-directed way (this is clear from looking at the criteria used to test for this), whereas ostensive communication involves the expression and recognition of intentions. Consequently, even if a species or group communicates intentionally, this does not mean that what is expressed are themselves intentions (which is what ostensive communication consists of, by definition). How, then, can we identify ostensive communication?

Since by definition it consists of the (1) expression and (2) recognition of both (a) communicative and (b) informative intentions, what we must look for is evidence of each of these four behaviours (i.e., 1a, 1b, 2a, and 2b). We have good experimental evidence for three of the four in human infants (the exception is the recognition of informative intentions), but the corresponding experiments in non-human species have not, to my knowledge, been conducted. It is therefore possible that other species do communicate ostensively, but my best guess, which I defend in this chapter, is that this is unlikely. It is more likely that ostensive communication is uniquely human. Why did it evolve in humans and only humans? My suggestion is that, since it critically depends upon recursive mindreading abilities, ostensive communication is a byproduct of increased skills of social cognition, which were selected for as an adaptation to the extremely social nature of human life. Once ostensive communication had emerged, it likely led to the selection of traits which allow ostensive communication to occur more easily than it otherwise would. An example is white sclera (the whites of the eyes), which allow us to more

easily infer the direction of eye gaze, which is an important means of ostensive expression. Natural pedagogy may be another such adaptation.

Chapter 5: Building a Language

As I noted in the opening remarks above, ostensive communication is a powerful tool, but its expressive potential is increased hugely by the creation of communicative conventions. How, then, do conventions emerge and become linguistic? Over the past ten or so years a literature has emerged in which these questions are studied in the laboratory. A variety of different experimental paradigms have been developed. Individuals motivated to communicate typically develop conventions quickly, although not without hiccups, and it seems likely that the same was true for the first "language" users. Over time, as these conventions are used repeatedly by the same individuals, or as new individuals acquire them, the sorts of structural properties that makes a system recognisably linguistic begin to emerge. A similar process has been observed in naturalistic settings, such as in the recent emergence of new sign languages in Nicaragua and elsewhere. These findings, and others, have been used to motivate a partial alternative to historically influential explanations of language structure. According to the traditional view, language structure is the consequence of a specialised faculty of language that constrains language form, and which allows humans to acquire and use languages. The alternative hypothesis is that language structure is a consequence of repeated instances of use and acquisition. This turns the previous view on its head: languages fit language users, not the other way around.

I believe this finding is best understood within the broad framework of Cultural Attraction Theory, which argues that, in the process of propagation, mental representations (e.g., beliefs, knowledge, intentions) and their public expressions (e.g., words, behaviours, artefacts) are non-randomly modified, in the direction of a better fit with the goals and dispositions of the human mind. They thus gravitate towards particular forms, called attractors, and away from others. This is attraction in its technical sense, taken from dynamical systems theory. Identification of the factors that influence this process can provide genuinely casual explanations of why cultural items take the forms that they do. The field in which this perspective has been most fruitfully applied to date is the cognitive science of religion, but this is certainly not the only example, and there is much potential to apply it more broadly. In the case of languages, models and experiments, such as those mentioned above, together suggest that two especially important factors are learnability (languages must be learnable by new users) and expressivity (languages must be able to express the meanings their speakers want them to). *Speaking Our Minds* presents these findings in terms of cultural attraction, but I probably could and should have stressed more clearly that by framing language evolution in this way, we can link explanations of structure in language with explanations of structure in other cultural domains.

Chapter 6: Evolutionary Adaptation

Chapter six aims to provide some answers to the following question: How should an adaptationist think about human communication and language?

To answer this question, it is important to keep in mind what the natural objects of study are here. Linguistic communication is not such a natural object of study, since there is no natural dividing line between it and other forms of ostensive communication. What might be a natural object of study is an innate cognitive mechanism – sometimes called a Universal Grammar – without which we would not be able to acquire and use languages. I say that this only "might be" a natural object of study simply because whether such a mechanism actually exists is a disputed and much vexed issue, on which I am personally agnostic. However, if there is such a mechanism, which shows signs of design for language acquisition and use, then the Darwinian conclusion that it is a biological adaptation seems inescapable. Such a mechanism could have been selected for once the use of communicative conventions, themselves built to enhance the efficacy of ostensive communication, has became widespread within a population.

Although linguistic communication is not a natural object of study, ostensive communication is such an object, and its ultimate evolutionary function is social navigation. Put simply, speakers speak in order to mentally manipulate their audience, and audiences listen in order to gain access to the minds of speakers. These direct functions can of course be subdivided into numerous derived functions, such as gossip, courtship, hunting, and all the other ends we use language and communication for. (It is important not to conflate direct and derived functions, which are two distinct levels of analysis.)

Finally, an important adaptationist question is what keeps human ostensive communication evolutionarily stable. My answer is the banal one that there are social costs to dishonesty that often outweigh the potential benefits. If we lie too often or in a too cynical way, our reputation will lead to rejection in future collaborations – a potentially severe cost for a species as social as humans. Some more elaborate explanations have been proposed, but these are typically based on misunderstandings of the underlying evolutionary theory.

Epilogue: The Big Questions Answered

The final chapter is very short. It lists the nine questions that leading researchers have proposed as the most important for language evolution. These questions include, for instance, "Why do only humans have language?", and "Where does language structure come from?". I then summarise how, unlike any previous work, *Speaking Our Minds* has answered all of them.

Comments

***

Hal Morris: Simulation Theory, Free-riders, Narrative

I am a layman who came to *Speaking Our Minds* via a sort of reading trek – from a Ray Jackendoff *Foundations of Language* (Jackendoff, 2002) reading group moderated by an anti-nativist who seemed to lose interest to Tomasello's *Origins of Human Communication* (Tomasello, 2008) (one good thing that came from the reading group).

A few months ago, Sperber and Scott-Phillips visited Scientia Salon, and got a dose of anti-nativism and anti-Cultural-Evolutionism (IMO). This piqued my curiosity about the two authors (and a third who stayed out of the SciSal discussion) and I read *Relevance: Communication and Cognition* (Sperber & Wilson, 1986/1995) and *Speaking our Minds*, and discovered this site.

I had also read Alvin Goldman's *Simulating Minds* (Goldman, 2006) a couple of years ago, which cited Tomasello's and his colleagues' evidence of fairly advanced mindreading at 14 months. The Simulation Theory of mindreading (distinguished by Goldman from Theory Theory or folk psychology) is not brought up in *SOM*, though I think it would be helpful. It seems overwhelmingly convincing to me. Goldman doesn't mention dreaming, but when we dream, we seem to be interacting with simulations of people in our lives, and it seems the most natural way to conceive of 3 or 4 level recursive mindreading coming naturally to a 14-month-old.

I would go so far as to wonder if our minds establish simulations of important other people as a permanent on-call bit of neural mechanism, like characters in the mind of an author, who, as some authors say, wait for the characters to tell her what to write next. Multilevel recursive mindreading would then be a rapid interaction between core self and simulated other, not a plodding inner conversation ("So she thinks that I think that she thinks...") – and the mechanism behind our spontaneous and sometimes irrepressible replays of frustrating conversations.

Assuming ostension based on *some* kind of mindreading, I was wondering what the debate about springboks and butterflies and cues, etc. was all about. I can only imagine that this has to do with rigorously justifying the dismissal of the "code model".

Is the free-rider discussion more of the same? I am having trouble engaging with that. The most likely context I can imagine for it involves the domain of language strictly for the purpose of reporting reality – that if all truly share experience, each one in a band, or the extended group of culture-sharers, get the benefit of the observations of all, that is a huge group benefit? But how do you hide being a free-rider? You need to say something. Make it up? But manipulative lying takes very high order cognition not automatically conferred by multilevel mindreading – especially if facilitated by neural features. Maybe early linguistic hominids just couldn't master it.

While many philosophers and generative linguists seem too exclusively focused on language as descriptive of reality, the examples in *SOM*, and especially the Peter–Mary examples, serve so well to illustrate ostension and mindreading as central to language (e.g., the examples about eating the berries, chopping the vegetables, and refilling the glass). Such exchanges seem indeed like they might be the subset of language most likely to evolve directly from primate gestural communication.

But how does narrative fit in? Mindreading of the audience seems minimal. Perhaps the most basic level of narrative is more or less direct translation of a linearized recollection of experience into words – only possible after considerable prior evolution of language, perhaps well after the solicitation of cooperation. Such an adaptation might have a huge impact, giving the band a sort of supermind with combined knowledge.

2

WHY DO CHILDREN BUT NOT APES ACQUIRE LANGUAGE?

By Richard Moore

For reasons of length I could not include more substantive objections to chapters three and four of Thom's book in my review of the book in the TLS (Moore, 2015d). However, since the gaps in his argument undermine his claim to have explained why humans but not apes acquired language, I don't think the issues are trivial. While I develop some of these points elsewhere [1], since some of those papers are not yet out, this discussion seems like as good a place as any to point out the reservations that I have with some of the central claims of the book.

Central to Scott-Phillips's explanation of why humans alone evolved language are, he claims, two cognitive abilities: the ability to act with and understand fourth order metarepresentations, and the ability to distinguish between 'informative' and 'communicative' intentions. Scott-Phillips argues that humans but not apes acquired language because they possess both of these abilities in ways that apes do not. I don't think this claim is defended adequately in his book.

As I mentioned in the review, there is currently no evidence that pre-verbal infants could understand fourth order metarepresentations. Indeed, there is evidence – not discussed in the book – that some six-year-olds struggle to entertain even second order metarepresentations (Perner & Wimmer, 1985). Scott-Phillips appeals to the existence of O'Grady et al. (2015) to show that human adults can entertain seventh order metarepresentations. However, since this ability may be a function of their language acquisition and not a prerequisite of it, and since Scott-Phillips's claim must be true of pre-verbal children if it is to support his conclusions, the O'Grady study just doesn't tell us what we would need to know.

This takes us to the discussion of the significance of the informative–communicative distinction. According to Scott-Phillips's view, ostensive-inferential communicators (including pre-verbal children) must be capable of each of the following:

(1) the expression of informative intentions

(2) the recognition of informative intentions

(3) the recognition of communicative intentions

(4) the expression of communicative intentions

I won't say anything about (1), (2) and (3) here. (There is more to be said about each one, but that may require a great deal of further unpacking. For example, there now exist data for (3) that show inconsistent roles for ostensive cues in both human and ape gesture interpretation; and while I agree with Thom's empirical claims about (1), I would have liked to see more theoretical support for the benchmark that he sets.) However, I do want to raise some issues with this discussion of (4), understanding of which Thom takes to be manifested in engaging in hidden authorship. I found his handling of this point to be disingenuous.

On Scott-Phillips's view, which he repeats in his 2015 *Current Anthropology* paper (Scott-Phillips, 2015b), it is a prerequisite of grasping the informative–communicative distinction that one should be able to act with hidden authorship, since that requires recognising the significance of the ability to inhibit one's communicative intention.

He supports the claim that pre-verbal children could distinguish between informative and communicative intentions by pointing to his own hidden authorship study (Grosse et al., 2013) on children. Although he refers to the subjects in this study only as 'children', the youngest children tested in this study were three years old, and so the study does not show that pre-verbal children are capable of hidden authorship in the way that his argument requires. It is not appropriate to draw conclusions about what pre-verbal children can do based on the abilities of three-year-olds. Indeed, when he and Gerlind Grosse ran their hidden authorship study in Leipzig at the end of 2009, they presumably also did not expect that pre-verbal children would show hidden authorship in their paradigm – since they did not attempt to test any children younger than three. It is a problem throughout Thom's book that he is inattentive to the ages of subjects when making claims about human ontogeny.

The fact that Scott-Phillips over-interprets his own hidden-authorship data might be more forgivable if he did not also claim that the absence of evidence of apes performing well in the same paradigm should be interpreted as an 'implicit collective acknowledgment' by primate researchers that apes would fail the task. I agree with him that in this case, they probably would. However, since we don't even know whether pre-verbal human children would succeed in this paradigm, it's hard to know how to interpret this. I am sceptical that they would do so – but I also think this inconsequential, since hidden authorship is the wrong marker to use (Moore, under review). Moreover, I am troubled that Scott-Phillips interprets an absence of evidence favourably in humans but unfavourably in apes, despite having reasons to doubt that the finding would really be present in pre-verbal children.

Since he thinks these abilities necessary for ostensive-inferential communication, and since his attributions of the relevant abilities to pre-verbal children remains empirically unsupported and tendentious, we ought to be sceptical of his conclusion about why humans alone evolved language. There may be many reasons why apes did not acquire language. My preferred explanation is that, unlike humans, they simply failed to grasp the coordinative potential of their existing communicative abilities – perhaps because they did not face the same ecological challenges as our ancestors, which forced them to overcome more challenging tests of coordination. I suspect that many of the abilities that he thinks are pre-requisites of ostensive-inferential communication are really acquired only after language. (I defend this view at length in the submitted paper mentioned in the footnote.) Even if my view is false, though, I don't think Scott-Phillips's discussion is sufficiently detailed to warrant any strong conclusions about the origins of language.

To give another example of his too quick treatments of important issues, there is no question that Michael Tomasello's hypothesis about the cooperative foundation of human language is a hugely important one. However, Scott-Phillips dismisses it in a few paragraphs at the end of chapter three, following only a very superficial discussion of the relevant literature on joint action. In fact, Tomasello's claims about the cooperative foundations of communication follow from his acceptance of claims that Thom also accepts – namely, the need for pragmatic interpretation of utterances. On Tomasello's view, pragmatic interpretation is possible only because speaker and hearer together take themselves to be engaged in a joint project of bringing the hearer to grasp the speaker's communicative goal. Scott-Phillips's cursory treatment of this point fails to get to grips with Tomasello's underlying (and admittedly not entirely clear) motivation, and so does not make it clear why Tomasello's view should not be accepted.

I hope that these comments won't be interpreted as showing hostility to Scott-Phillips's work. While I disagree with much of what he says, he is a wonderfully clear thinker, and my own ways of thinking about these issues have benefitted hugely from the perspicacious ways in which he carves up the conceptual terrain. At the same time, his preference for simple, clear answers leads him to ignore details. While it is too much to expect a book of 200-odd pages to answer all of the problems that it raises, it is not unreasonable to expect it to temper its conclusions in the absence of attention to detail. Scott-Phillips's claim to have explained why humans alone evolved language is unsupported by the arguments that he offers.

On a general level, it may be that given Thom's fondness for big pictures over details, he won't worry about these complaints. As he says in his book:

> "There is nothing wrong with lumping in science. On the contrary, it is how we develop our major theories and paradigms."

The problem with this view is that ultimately it is details that will falsify our major theories. Ignoring them is therefore not an option.

[1] I develop these points further in three related papers – my response to his *Current Anthropology* paper (Moore, 2015a), my response (Moore, 2015e) to his recent *Animal Cognition* paper (Scott-Phillips, 2015a) and an original paper (Moore, 2015c). In both of the response papers I spell out an alternative test of the informative–communicative distinction. In the cognitive development paper I argue that neither fourth order metarepresentations nor sophisticated folk psychological abilities are a prerequisite of ostensive-inferential communication. If anyone is interested, all of these can be downloaded from my page on ResearchGate.

Comments

***

Thom Scott-Phillips: A defence against some serious charges

In these critical comments, Richard takes issue with one of the central claims of *SOM*, namely that humans but not any other great apes engage in ostensive communication. A full discussion of the issues he raises would require a response that is longer and more detailed than I can give right now, and so here I focus only on defending *SOM* from some of his more serious charges.

First, some context. I defend in *SOM* the claim, made by others before me, that engagement in ostensive-inferential communication proper depends upon the competent use of recursively-embedded mental states (see in particular §3.4). There is a tension between this theoretical position, on the one hand, and a common sense intuition, on the other, that processing recursively-embedded mental states is cognitively difficult, and certainly not something that can plausibly be involved in everyday communicative interactions. Something, then, has to give. Either accept the theory and reject the intuition, or vice versa. As I mentioned in my précis to this book club "most researchers who have grappled with this problem have chosen to accept the intuition, or some version of it, and argue against the theory". Richard is one such researcher, and he has made a number of substantial contributions to this discussion. In *SOM* I take the opposite position – I argue in favour of the theory and against the intuition – and it is this view that Richard takes issue with, both here and elsewhere. Although there are several important points of disagreement between us, I would like to stress that I am grateful for Richard's detailed critique of my work, since he rightly brings attention to several points that do require further elaboration and defence.

In particular, Richard is correct to emphasise that we do not (yet) have the evidence necessary to fully accept some of my claims. For instance, Richard questions whether young children, who are competent users of ostensive communication but who have not yet acquired language, are able to process the sort of recursively-embedded mental states that, I argue, are in fact involved in ostensive communication. We do not know at present whether this is true.

However, I was not concerned in *SOM* with establishing this claim, and related ones, beyond doubt. My agenda was more modestly to argue for its plausibility. To this end, I reviewed in chapter three evidence that even pre-linguistic children engage in mindreading, and briefly mentioned one adult experiment that suggests that recursively-embedded mental states pose no particular cognitive challenges for adults. There is, as Richard points out, no way that such data could ever establish that my claims are fact. Such data can however establish that my claims have some degree of plausibility. I also, moreover, presented arguments that my claims are more plausible than alternative positions. This led me to conclude that "the idea that both adults and children's communication really does involve recursive mindreading is not only plausible, but is in fact the most parsimonious interpretation of the data at the time of writing" (p. 74). Richard's comments give the impression that I believe that these matters are fully resolved, but as I hope this quote illustrates, I do not believe this, and *SOM* makes no such claims.

Subtle misreadings of my views recur throughout Richard's comments. Regarding hidden authorship, Richard writes that "On Scott-Phillips's view [...] it is a prerequisite of grasping the informative–communicative distinction that one should be able to act with hidden authorship". Yet what I actually wrote is that command of hidden authorship "comprises good evidence of an understanding of what a communicative intention consists of, and its relationship with informative intentions" (p. 89, italics added). I hope I am not being pedantic here. Good evidence is not the same thing as a prerequisite. Indeed, in the present context one is a fair presentation of my view, but the other is not.

Still, this is incidental to the main point that Richard wishes to make with hidden authorship, namely that *SOM* does not pay sufficient attention to the exact ages at which children pass the various social-cognitive tasks under discussion. There is some partial truth to this. It is at least possible, as Richard suggests, that children in fact pass these tasks only after they acquire language. Moreover, if this is true, it would undermine my claims, and it is probably fair to say that *SOM* should have made this point more explicitly.

What is not fair, however, is to fault a book for not being a different book. Academic output should be critiqued on its own terms. Richard believes that *SOM* should have paid more attention to several other claims – he offers Mike Tomasello's work as an example – but it was never the agenda of SOM to provide a detailed review of the entire terrain. On the contrary, in fact. Here is what I say in the Preface to *SOM*: "I have tried, as much as possible, to make my arguments positive ones, in favour of a particular view of the origins of language. I am of course critical of other perspectives where it is necessary or useful to be so, but in some cases I have avoided direct confrontation with some views that differ from my own, in order to maintain a focus on the positive case for my own views." (p. xiv). I make related remarks at several other places, including just before my discussion of Tomasello's work (p. 75). I see this focus on making a positive contribution as a strength rather than a weakness of *SOM*, but either way, these are the terms on which *SOM* should be judged.

It is unfortunate that Richard's comments end with an objection that comes close to ad hominem criticism: "it may be that given Thom's fondness for big pictures over details, he won't worry about these complaints". The irony is that to support this assertion, he ignores detail himself: he quotes SOM directly, but out of context. The quote he uses – "There is nothing wrong with lumping in science. On the contrary, it is how we develop our major theories and paradigms" – comes from §3.3, where I defend the claim that Relevance Theory provides a proper scientific paradigm for pragmatics. As Richard knows, the development of new scientific paradigms is all about the search for generality in one's explanations, and it is in this context alone that I argue in favour of lumping. I certainly not do not make any arguments in favour of ignoring details.

Richard Moore: A brief rejoinder

I appreciate Thom's comments and I am happy to concede some of them. It is true that in the case of hidden authorship he talks about 'evidence for' rather than 'necessary for' – and I misrepresented him to the extent that I didn't make that clearer. At the same time, it is still not true that evidence of a certain behaviour in three-year-olds is evidence of anything in pre-verbal children; and so it is not clear that three-year-olds' ability to engage in hidden authorship can tell us much about why humans but not apes acquire(d) language.

I also appreciate Thom's point about the scope of his book. However, as Olivier notes, it is still a book that claims to have solved every major problem in the study of language evolution. Since this is a somewhat immodest claim, I don't think it is inappropriate to hold someone who makes it to a robust burden of proof! I am also genuinely puzzled by his presentation of his view as unorthodox (for example, in the sentence he quotes from his précis). In fact, versions of Thom's view have been defended by Michael Corballis, Sperber and Wilson, and – most extensively – Tomasello. By contrast, the 'standard view' re-interpretations of Grice have come in a neglected book chapter by Juan Carlos Gomez, and mostly unpublished work by me. What Thom adds to the Tomasello's and his colleagues' view – and does with great clarity and nuance – is a clear acknowledgement of the precise cognitive abili-

ties that they implicate; and he replaces Tomasello's evidence on the need for cooperation with an appeal to the explanatory mechanisms posited by Relevance Theory.

As for the final comment, I didn't intend for it to come across as a dig, and I am sorry if that wasn't obvious. The text I quoted was intended just to be a reflection on the differences in Thom's approach and mine to the questions that interest us: we have had the splitters vs. lumpers argument many times over, and we fall on very different sides. Of course, I don't think Thom has ever argued that differences should be ignored – even if I think he is sometimes insufficiently attentive to them.

3

A FEW COMMENTS ON 'SPEAKING OUR MINDS'

By Paulo Sousa and Karolina Prochownik

We would like first to thank the ICCI team for the invitation to participate in the book club around Thom Scott-Phillips' *Speaking Our Minds*, where a new theory of the evolution of human language and communication is put forward. This is a fantastic book full of groundbreaking ideas, and we are pretty much persuaded by the proposal. We are not experts on the topic and in our commentary we shall focus on some aspects of the book that we thought could have been approached from a slightly different angle or could have been a bit more elaborated, without presuming that our concerns challenge the core tenets of the book in any respect.

Communication

Scott-Phillips characterizes communication as a type of interaction that involves actions and reactions that have functional interdependence, where the source of the function may be natural selection or human intentionality (p. 29–33). We couldn't avoid the feeling that this characterization is not specific enough. Take a physical fight between members of two species with attack–defense dispositions that evolved complementarily (e.g., in the context of a predator–prey relationship). All behavioural aspects of the fight seem to be communicative according to the above characterization, but it seems to us that only some aspects related to the transmission of information are communicative, and the above characterization does not seem to have the conceptual resources to specify the relevant aspects.

Scott-Phillips also uses the label "signal" instead of "action" in the above characterization, which may suggest a way of constraining the above characterization to include only actions that constitute signals or to refer only to the signalling aspects of actions. However, the meaning of the label "signal" in the context of the above characterization seems to be derived completely from the notion of functional interdependence (see p. 30, Table 2.1), and other characterizations of "signal" given in the book do not seem to entail a more specific characterization either (p. 25–26). Scott-Phillips wants to prioritize the notion of functional interdependence and neglect the notion of information transfer in the cha-

racterization of communication (p. 32), but perhaps only with the inclusion of some element of information transfer in the characterization, can one avoid our apparent problem.

Ostensive communication

Sometimes Scott-Phillips explicitly emphasizes that the property of overtness in signalling is fundamental to the characterization of ostensive communication: "…ostensive-inferential communication can be glossed as *intentionally overt communication*" (p. 23); sometimes he is less explicit in this respect (p. 10–13). Perhaps this oscillation mirrors a slight inconsistency in the discussion. On the one hand, Scott-Phillips seems to argue that only with overtness there is the property of functional interdependence that characterises communication in the human context – e.g., when stating that, among various scenarios where Mary transmit to Peter the information that the berries are edible, there is human communication only when she overtly transmits the information (p. 66–67). On the other hand, he seems to suggest that overtness is not a fundamental condition for functional interdependence/communication in the human context – e.g., when classifying hidden authorship, which does not involve overtness, as a *type of communication* (p. 89, Table 4.1). In our view, a more focused and detailed discussion of the notion of functional interdependence in the human context, dealing with questions such as whether and why overtness is (or is not) a fundamental condition for communication, is missing in the book.

Combinatorial communication

Scott-Phillips states that there are three types of routes to combinatorial communication (or to the creation of composite signals) – ritualization (from cues), sensory manipulation (from coercion), and the direct route (from ostension). A substantial part of chapter two of the book is devoted to the discussion of combinatorial communication in natural codes in contrast with combinatorial communication in human language. The general idea is that combinatorial communication is poor in natural codes because it can only emerge by ritualization and sensory manipulation, while it is rich in conventional codes because it arises via the direct route of ostension. However, we think that the discussion actually conveys two different pictures about the existence of combinatorial communication in natural codes.

For the most part of the discussion, Scott-Phillips claims that one could find some real (though simple) cases of composite signals in natural codes, and that examples of comparable complexity may occur across very different species (e.g., pyow-hack calls in putty-nosed monkeys and combined molecules in the bacterium *Pseudomonas aeruginosa*). According to Scott-Phillips, this real but rare existence of composite signals in natural codes across species is to be explained by the fact that the emergence of composite signals via a process of ritualization or sensory manipulation has a quite low probability, and a probability that is independent of the variation in cognitive capacities across species:

> "The rarity of combinatorial communication in the natural world cannot be explained by cognitive differences between species, since bacteria have a very different 'cognitive' system to our own, yet they seem to have a communication system that, in terms of its combinatorial complexity is as close to human language as any other." (p. 50)

However, in another short but relevant passage, Scott-Phillips suggests a different picture:

> "There is a sense in which natural codes like this bacterial one are not really combinatorial at all. After all, there is no "combining" going on. There is really just a third holistic signal, that happens to be comprised of the same pieces of existing holistic signals. The same seems to be true of the putty-nosed monkey calls: 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. In contrast, the ostensive creation of new composite signals is clearly combinatorial: although the meaning of the composite signal is not the exact sum of the meanings of the component parts, it is in part a function of those meanings (and the context)." (p. 50)

In other words, here the idea is that the processes of ritualization and sensory stimulation do not lead to a composite signal in any relevant sense, which would entail that in fact the only real route to combinatorial communication is the direct route of ostension.

In our view, the second picture, which was much less emphasized, implies a kind of explanation that is even more consistent with the overall thesis of the book that human language is qualitatively different from a natural code.

Communication in apes

There were two aspects of Scott-Phillips' discussion of communication in apes that we thought could have been more elaborated. The first aspect is related to a brief passage concerning attempts to teach apes some forms of non-vocal language: "…this approach had some success, in the sense that at least some of the apes developed competent command of communicative systems that have some, although certainly not all, of the distinctive features of human languages" (p. 80). It seems to us that a more detailed analysis of this acquired competence could have shed more light on the debate concerning the communicative capacities of apes.

The second aspect concerns Scott-Phillips' main claim that natural codes in apes are quite flexible, and that this is due to apes' incipient metapsychological capacities: "What we appear to have, then,

is a system made possible by mechanisms of association, and made expressively more powerful by the existence of metapsychological capacities, which allow a natural code to be used in a flexible way." (p. 93). Although the author devotes more space to this issue in the book, no vivid illustration of the flexibility in coded communication among apes is given; so, as non-experts on the topic, we couldn't grasp what exactly is involved in this flexibility. Moreover, although Scott-Phillips provides a detailed characterization of apes' incipient metapsychological capacities, he does not discuss in detail the casual connection between these capacities and flexibility in coded communication – we couldn't grasp what exactly is involved in a code being made expressively powerful by incipient metapsychological capacities. On the side of the interpreter, for instance, is it that the incipient's metapsychological capacities enhance some probabilistic decoding? Or is it perhaps that the incipient metapsychological capacities introduce some inferential ability utilized to interpret flexible but purely semiotic messages?

To conclude, we would like to reiterate that we don't take our points to undermine in any respect the great contribution of Thom Scott-Phillips' *Speaking Our Minds* to the debate on the evolution of language and communication. And we look forward to the further developments of his exciting research program.

Comments

***

Thom Scott-Phillips: Some conceptual clarifications, and avenues for elaboration

I'd like to begin this response by thanking Paulo and Karolina, as non-specialists, for taking the time to read SOM in detail, and to write their thoughts down for us all to read. Anybody who has tried to communicate across disciplinary boundaries will appreciate that there is always much potential for misunderstanding. Thus, by providing these comments, Paulo and Karolina give me some insight into what parts of SOM are most comprehensible and persuasive, and which are less so, and this alone is reason for me to be grateful to them. But more than that, they also make some positive suggestions about which of the concepts that I introduce and/or make use of in SOM could and should be developed and elaborated further.

Paulo and Karolina divide their comments into four parts, and my response will mirror this structure.

Defining Communication

Paulo and Karolina make the same suggestion that Dan did in his commentary, namely that my definitions of communication and associated terms may not be specific enough. This confirms to me that SOM does not express clearly enough an important nuance of my definitions, namely that the action and reaction are not just designed, but designed "for the purposes of playing that role [action or reaction]". This is exactly the reason why I talk about communication in terms of functional interdependence. I believe that once this nuance is added, it addresses the issues that Paulo and Karolina raise. I should admit however that I have always had trouble expressing this nuance. It is there in my 2008 paper on this topic (Scott-Phillips, 2008), but the language is cumbersome, and hard to follow. As I said in my response to his comments, Dan's suggestion that the notions of proper and actual domains could be useful here is probably right.

Is overt intentionality part of communication?

Here too Paulo and Karolina's comments pick up on subtle but important conceptual distinctions, and force me to be clearer. Let me address their query directly: I believe that functional interdependence is a defining feature of communication. Paulo and Karolina are correct however that some passages of SOM suggest otherwise. In particular, I imply that hidden authorship, which does not involve functional interdependence, is an instance of communication. This was a mistake: hidden authorship is not strictly communicative (precisely because it does not involve functional interdependence). In short, the reason hidden authorship is interesting is not because it is communicative per se, but rather because, despite not being strictly communicative, it demonstrates a command of important cognitive aspects of human communication.

Combinatorial communication

Paulo and Karolina bring attention to an observation that SOM makes only in passing, namely that where we do see combinatorial communication systems in nature, this is still not "combinatorial" in the relevant way. Paulo and Karolina point out that SOM does not much elaborate on what this relevant way is. Let me therefore make a short comment to that end here. In §1.5 of SOM I distinguish between natural codes and conventional codes. My point in the passage that Paulo and Karolina quote is that even where we do see combinatorial communication in nature, it is still a natural code – but when we see it in linguistic communication it is a conventional code. Consequently, even if natural codes do develop combinatorial signals, the ontology is still different to the linguistic code, and that limits the inferences that we can make with evolutionary comparisons between the two.

Great ape communication

I agree with Paulo and Karolina suggestion that there is significant scope for elaboration of my characterisation of great ape communication. They write that "as non-experts…we couldn't grasp what exactly is involved [in great ape communication]". They should not worry: neither do I, as the proposer of this idea, know the details! Put another way, I agree that if one wishes to understand great ape communication for its own sake, then my suggestion that it is best characterised in terms of an intentional code is really only a starting point. I did not pursue it further only because it is not central to the thesis of SOM, and I wished to keep the book slim and focused. I would be delighted to see my proposal elaborated on, and Paulo and Karolina are right to suggest that the impact that teaching can have on great ape communication will be critical to any such elaboration.

4

KEY NOTIONS IN THE STUDY OF COMMUNICATION

By Dan Sperber

I am enthusiastic about Thom Scott-Phillips' book. It integrates cutting-edge research in several fields, from biology to pragmatics, relevant to the study of the evolution of human communication and it redirects the whole enterprise in a new, much more promising direction. This, however, is not the place to wax lyrical about the book; so, let me focus on a single conceptual question of broad relevance where Thom makes a very valuable contribution, but one that, I believe, still needs more work on the part of all of us.

In chapter two, Table 2.1, Thom proposes a set of "definitions of key terms in communication" developing the idea that "if an action [of one organism] causes a reaction [of another organism], and both are designed for the purposes of playing that role in the interaction [between the two organisms], then we can term the action a signal and the reaction a response, and the overall interaction communicative. If, however, one half or the other was not designed for these purposes, then the action is either a cue or a coercive behaviour."

	Action designed?	Reaction designed?
Signal/Response	Yes	Yes
Cue	No	Yes
Coercion	Yes	No
Accident	No	No

The idea of defining key notions in the study of communication and related phenomena in term of the presence or absence of interfacing designs or functions is a brilliant one (suggested by John Maynard Smith and David Harper and developed in an original way by Thom), but one that raises a number of difficult issues. I'll focus on just one such issue.

The issue comes, in good part, from the fact that this approach is rooted in ethological approaches to communication. There, what matters is the behaviour of the interacting organisms. Their internal cog-

nitive states and processes, if considered at all, are just a part of the mechanism causing the expected behaviour. What matters is the behaviour.

The ethological approach, whatever its great merits, isn't such a good source of insight when you study forms of communication the success of which cannot be identified in terms of any behavioural reaction, as in much of human communication. The basic reaction to communication is, in all cases, cognitive; it may, as generally in animal communication, be a trigger to a specific behavioural response. Even in the human case, the cognitive reaction may be aimed at causing a specific behavioural reaction, as when what is communicated is a request ("Pass the sugar, please!"). Much of human communication, however, isn't aimed at causing some specific behavioural response. The production of this post, for example, is an instance of communicative action designed to elicit in you, readers, just a cognitive reaction. This post might also elicit a behavioural reaction is some of you (yawning or buying Thom's book, for example), but this is not what it is designed to do.

Thom's definitions are more appropriate for communication aimed at a behavioural response. Even there, the cognitive component makes communication special, in a way that is not reflected in his definitions. Thom's definitions, actually, would be appropriate to talk not just of communication but of all forms of interaction where there is a difference between action and reaction. Communication, however, is about providing a cognitive input and causing a cognitive reaction (which may but need not be a way of in turn causing a behavioural reaction). I would suggest that this should be reflected in the way it gets defined. Otherwise, we extend our notion in a way that I find unhelpful.

With Thom's definitions, somewhat paradoxically:

(1) Mating would provide in itself the perfect example of communication since the action of one organism and the reaction of the other are each designed for this interaction. This is true not only of mating in animals that commonly do communicate in, around, and through mating, but also of mating between, say, two haploid yeasts of opposite mating types the behaviour of which can be parsimoniously described without invoking at all the notion of communication.

(2) The relationship between beetles feeding on elephant dung and elephants would count as one of 'cue': the dung being the cue that the elephants produce and from which beetles benefit, the production of the benefit being an undesigned action and the feeding being a designed reaction.

(3) Any designed type of action of one organism on another with no designed reaction (for example the action of a parasite plant feeding on its host) should count as "coercion" – which it is, of course, in this example but in the ordinary sense of 'coercion' rather than the one intended by Thom. (Actually, Thom's example in his book of a woman being pushed from her chair is a case of coercion in the ordinary sense, so, I may be misunderstanding him.)

The word "coercion", by the way, works much better for non-cognitive interactions than for cognitive ones. For the cognitive case, I much prefer "manipulation" for what Thom has (or should have) in

mind. "Coercion" may be a well-established term in one strain of biological writing on the issue, but "manipulation" has a good pedigree too, including in biology, and well beyond.

Why, you might well ask, not use for communication a classification that works for all kind of interactions? Because communication has some quite specific features that would not be properly identified under such a general definition. There is, in particular, a basic asymmetry of major evolutionary significance between the "reaction", which is primary, can stand on its own, and doesn't need any manipulative or signalling "action" to be adaptive, and the "action" side, which is adaptive only if there are receivers with a cognitive capacity to react.

Take the "cue" situation. It is just the ordinary situation of any organism involved in any perception-based cognitive activity. Such an organism has mechanisms designed to take advantage of some specific features in its environment, features that indicate something of relevance to it and that are therefore cues (in the ordinary sense at least). Cues in this sense may be provided not only by other organisms but also by anything that can affect sensory receptors. Organisms pick cues of impeding changes of weather, cues of the presence in useful quantity of various important substances, water, minerals, and so on. Cues of biological origin are, of course, centrally important: much of cognition is in the business of recognising and processing cues of the presence of mates, or of prey, or of predators, i.e., other organisms (which also have an interest in being or not being recognized).

Insect-eating birds are equipped to detect insects, and more specifically, to discriminate types of insects according to the benefits and costs there may be in eating them. The distinctive features of various species of insects provide cues for this discrimination. It may be in the interest of some species of insects to evolve (1) features that make them bad prey, and (2) features that advertise the fact that they are bad preys, as is famously the case with Monarch butterflies, which are poisonous and highly recognisable. The distinctive pattern of their wings is a classic example of interspecies "signalling", in Thom's exact sense: it is designed to cause a reaction in birds and this reaction is itself designed. My point here being that you need cognition at the reaction end of the interaction, but not necessarily at the action end: the Monarch butterflies' pattern is not an output of their cognitive processes. More generally, signalling, even in its broad biological sense, is an exploitation of cognitive capacities in the recipients. This signalling can be "designed" by natural selection. It needn't be intended or even cognised in any way by the signaller. Signalling needn't involve any cognitive capacity.

Monarch: signalling.
Viceroy: manipulating

Source: Wikipedia

Describing "coercion" or, as I prefer to call it, manipulation, as involving a designed action but no designed reaction is not strictly incorrect, but it may be quite misleading when one sticks to cases of cognitive reaction. Manipulation works by exploiting a cognitive mechanism, and it does so by providing an input with features that the mechanism is designed to react to but that, in this case, – and this is what makes it manipulation rather than signalling – provide the wrong input for further inference. So there is a designed reaction that is being exploited, but its function isn't to be so exploited, making the case different from that of signalling.

The Viceroy butterfly, whose wing pattern mimics that of the Monarch, is picked up by a cognitive mechanism in birds that evolved to pick such patterns – the mechanism isn't precise enough however to distinguish the Viceroy from the Monarch exact pattern (see the figure); you may have trouble doing so yourself –, but it leads in this case to the mistaken inference that the butterfly is poisonous and shouldn't be eaten (actually, as Olivier Morin warns me, whether the Viceroy is genuinely palatable and hence whether this whole classical example of "Batesian mimicry" is genuine has become controversial, but even so, it serves my illustrative purpose). Unlike coercion as defined by Thom, manipulation in this sense is restricted to actions that take advantage of the limits of the receiver's cognitive mechanism by producing, in the case of mimicry, false positives, and, in the case of camouflage, false negatives (I would describe all this in terms of the proper and actual domains of the cognitive mechanisms involved). In other terms, the possibility of coercion/manipulation is dependent on the existence of cue-picking mechanisms whereas the possibility of cue-picking (i.e., of cognition) isn't symmetrically dependent on the existence of coercion/manipulation.

One of the effects of this dependence of manipulation on genuine cues and, often, on actual honest signals, is that there is a continuum of cases between manipulation and signalling. Think of the expression of genuine emotions that can also involve a good dose of manipulation.

Thom's "definitions of key terms in communication" raise several other important issues having to do, for instance, with the relationship between biological and cultural functions (both well-exemplified in human language) and with the graded vs. categorical character of these or any such definition (as suggested by the signalling manipulation continuum I have just alluded to). So I tend to see these definitions as a good provisional tool for the use Thom makes of them, but also as a way to frame some very basic and not yet resolved conceptual issues in the study of communication.

Comments

***

Thom Scott-Phillips: Refining biological communication?

Before I respond to his comments, let me take this opportunity to express in a public forum my deep admiration for Dan's work, which has had a very substantial and profound effect on my thinking. This is true both of his long-term collaboration with Deirdre Wilson, and of many of his other projects. *Speaking Our Minds* would be a very different book if not for these prior contributions, and it is my sincere hope that readers are able to see and appreciate this. It is, moreover, a delight for me that in his comments above Dan reverses the trend, by taking some of my work on defining communication, and building on it himself.

As Dan observes, I have followed a profitable tradition, well-established in ethology and related disciplines, of defining communication in terms of functionality (rather than in terms of, say, mechanism). He goes on to suggest that while useful, my framework requires further refinement if it is to address some "very basic and not yet resolved conceptual issues in the study of communication". By way of illustration, Dan proposes that my notion of "coercion" might be better characterised as "manipulation", and that whatever we call this category, it may exist on a continuum with signalling proper.

To respond to this, let me first clarify how I think of the term "coercion". Table 2.1 of *SOM*, which Dan reprints above, effectively provides a gloss of my definition, but there is a nuance not captured by the table itself. In the table's legend (and in the main text), I write that what matters for my definitions is not simply that the actions and reactions in question are designed, but whether they are designed "for the purposes of playing that role [action or reaction]". This is cumbersome language, but I hope the substantive point is clear: reactions are coercive not only when there is no designed reaction of any sort (as in, say, the example of somebody pushing a work colleague off her chair; see Figure 2.2 of *SOM*), but also in those cases where there is a designed reaction, albeit not one that is designed to be a reaction to the action in question.

Dan proposes that the language of proper and actual domains would be useful here, and I suspect he is right. Replace in Table 2.1 the heading "Reaction designed?" with "Is the reaction designed, and does the action fall within both its proper and actual domains?" (a corresponding change could be made to the first column too, if desired). Signalling describes only those cases where the action falls within both the proper and actual domains of the reaction; and coercion describes all other cases. Then the sort of cases that Dan points to, such as the wings of the Viceroy butterfly, do fall properly within the definition of coercion, as they should. I suspect that this nuance also handles the otherwise paradoxical cases (1) and (3) from Dan's comments. (I am not entirely clear why case (2) is paradoxical. The presence of the dung seems to me to be a cue, and correspondingly satisfies the conditions for being so. Perhaps I have misunderstood the problem.)

Dan proposes instead that signalling and coercion (or, as he would prefer it, manipulation) exist on a continuum. I am not certain if I have read him correctly, but the argument I infer from his comments is that actions might fall within the actual domains of more than one mechanism, and hence there is a continuum of cases, from, at one end, those cases where the action falls within the proper domain of all the mechanisms involved (this is the signalling end of the continuum), and, at the other, those cases where the action falls only within the actual domains of the mechanisms, and not their proper domains (this is the coercion/manipulation end of the continuum). This is a very thought-provoking proposal, which strongly merits further development and reflection, not just from myself or Dan, but from all of us concerned with a cognitive approach to human communication.

Dan Sperber: A definitional Rubicon?

Thom is right to stress that, in his definitions of signalling, cue, and coercion, what matters "is not simply that the actions and reactions in question are designed, but whether they are designed for the purposes of playing that role [action or reaction]". I was well

aware of this point in my discussion and should have mentioned it explicitly. This however doesn't answer my concern: Thom's definitions are such that there is nothing to restrict them to the case of communication (and the related cases of 'cue' and 'coercion/manipulation'). They apply quite naturally to all interactions between organisms where an action triggers a reaction.

(1) If both action and reaction are designed to mesh with the other, then it is a case of mutual benefit between actor and reactor (mutually beneficial communication being a special case). (Incidentally, this rules out as an example falling under Thom's definition the "physical fight between members of two species with attack-defense dispositions that evolved complementarily" evoked by Paulo Sousa and Karolina Prochownik in their comment, since the attack isn't designed to trigger the defense: if the prey doesn't fight all the better for the predator).

(2) If the action is designed to trigger a certain reaction but the reaction isn't designed to react to that particular action, this is a case of the actor benefitting asymmetrically from the reactor, or 'coercion' (cognitive manipulation is a special case).

(3) If the action, without being designed to do so, does trigger a reaction designed to take advantage of the action, then this is a case of the reactor benefitting from the actor (benefitting from cognitive cues being a special case).

Note that, in the general case, the difference between cases (2) and (3) is a matter of who benefits, the actor or the reactor. Quite often, the actor–reactor distinction is itself a matter of descriptive choice: free-swimming algae arrive inside an oyster that eats them – is the actor the alga that moves into the oyster triggering its feeding reaction (and then it is a case of 'cue', in Thom's terminology) or is it a case of 'coercion', the oyster coercing the alga's movement and causing its decomposition? Both definitions work, which isn't as good as it sounds.

Now, in the case of cognition and communication, the asymmetry between the role of the source and the receiver of information is basic. I was suggesting that good definitions of communication and related phenomena should reflect this and therefore not just be a special case of general definitions of action–reaction pairings.

The point is illustrated by the case of the dung beetle that I used in my post as something that shouldn't be considered a cue. So, the elephant produces dung on which the beetle feed: benefit to the reactor in this case, 'cue' in Thom's sense. Where is the problem, Thom might say? A slightly more fine-grained description would distinguish two related reactions from the beetle, (1) a cognitive reaction: the beetle recognises the substance as food (and for this, properties of the dung that the beetle is equipped to perceive serve as cues), and (2) a more 'practical' reaction: the beetle eats and digests the dung; here to describe dung as a 'cue' seems to me unhelpful or even misleading. To make the point clearer, compare the beetle and a plant the growth of which also benefit from the elephant's dung. Is dung a 'cue' to the plant too?

The issue – or the stake – is a huge one: since Aristotle, Port-Royal grammarians, Peirce, Saussure, Goodman, and many others, insightful proposals have been made to define a number of notions including communication and around it. None of them is without problems. I believe Thom's approach may well be aimed in a promising, novel direction, but we are not there yet. Or should we go all the way to the Rubicon and just catch some fish?

5

CULTURAL ATTRACTION, "STANDARD" CULTURAL EVOLUTION, AND LANGUAGE

By Alberto Acerbi

Speaking Our Minds was a great pleasure to read. This slim book provides even a non-expert like myself with an accessible but, at the same time, in-depth treatment of language evolution. Scott-Phillips proposes us a coherent and, according to him, exhaustive, picture of the origins and evolution of language. The big questions are answered: we can proceed to the next topic.

I wonder how the community of linguists will feel in regard to this bold attitude (by the way, I am all for bold attitudes). As for myself, I can comment on a particular aspect of the book, that is, the role assigned to cultural attraction in explaining some of the features of language. The basic idea behind the concept of cultural attraction is spelt out with remarkable clarity in *SOM*. In short, cultural transmission, differently from biological transmission, is mainly a reconstructive process. Each time we "copy" a cultural trait we are in fact reconstructing it, starting from some piece of information that we gather from others. Individual modifications are not rare, and they are not errors. They are the crux of the cultural transmission process, and, importantly, they tend to be oriented in non-random ways (hence the notion of attractor).

One example – which I discovered reading this book – is tonal languages. In tonal languages (like Mandarin Chinese) the pitch that one uses to pronounce a word makes a difference to its meaning. It has been discovered that the distribution of tonal languages is associated with the distribution of two genes that regulate neural development. These are not genes *for* tonal languages, as individuals without the genes can learn them (and vice versa), but they may represent a factor of cultural attraction if these genes make it easier (for example) to detect or produce pitch differences. Imagine a population in which few individuals have the variants. Language changes that give more importance to pitch will be, in this population, rare and generally not reconstructed. The population will converge on a non-tonal language. The opposite will happen in a population in which the majority of individuals have the genetic variants in question.

Scott-Phillips gives a few other examples of factors of attraction that may shape language attributes,

some related to biological or cognitive features (like the example above), and others related to communicative needs, drawing mainly on the research from the Edinburgh Language Evolution Group. Overall, the case is convincing: cultural attraction is likely to have an important role in determining the features of the languages we speak today, and the details of their evolution. But is that all? What about all the researches that use a more "standard" evolutionary framework to study language, that is, that consider it like a culturally transmitted *replicator*?

The idea that languages evolve like biological species, with a process of descent with modification, has a long and successful history. Darwin's famous endorsement of language evolution testifies to that. Phylogenetic analyses are today used routinely in cultural evolution, and while their application to different domains is far from being uncontroversial, their success is at least partly due to the fact that they have been productively applied to language evolution, providing stimulating results. If phylogenetic analysis works for languages, what does this tell us about the feasibility of using standard evolutionary tools to understand their historical dynamics?

Recent researches showed that the rate of changes of words is correlated with their frequency of use. Words that are similar in related languages (a classic example is terms for numbers: think about *one* in English, *un* in French, *uno* in Italian, etc.) are also words that evolve at very slow rate, and, interestingly, are the words that are used with high frequency in daily life. This suggests a classic evolutionary pattern, one of generally faithful transmission with random modifications. Frequency of use would indeed affect rates of replacement by reducing the "mutation rate", as words used frequently would be, for example, remembered more easily than words only rarely used.

My general perspective is that various domains of human culture are characterised by various degrees of reconstruction and preservation in the transmission of their traits, and when domains are close to the "preservative extreme", it is useful, for pragmatic reasons, to consider them as standard evolutionary systems. Moreover, in the same cultural macro-domain, like language in this case, different aspects may be situated in different regions of the preservation/reconstruction continuum. More than asking which aspects are in general more important, it may be more productive to ask when and why transmission is preservative or reconstructive, and what the consequences are for the resulting cultural dynamics. For example, one may wonder whether the contemporary widespread use of media favouring strongly preservative transmission (such as "sharing" something on Facebook, or "re-tweeting" it) may play a role in contemporary language evolution.

In sum, I strongly believe that the cluster of ideas surrounding the notion of cultural attraction (the importance of individual reconstruction in cultural transmission, the fact that modifications to cultural items are generally not random, the importance of universal, or at least relatively stable, factors of attraction), developed in the past years by anthropologists like Dan Sperber and others, is one of the most important contribution to the contemporary study of cultural evolution. I am also open to considering whether cultural attraction forces are responsible for the most *interesting* attributes of languages, as one could infer from Scott-Phillips' book. A further step would be to identify which features of lan

guages are due to cultural attraction forces and which features are due to processes included in "standard" cultural evolution models, such as random modification of words, simple contextual learning biases, and similar, and how the various processes interact. The material presented in *Speaking Our Minds* may be an excellent starting point for this endeavour.

Some references…

On cultural attraction: see Claidière et al. 2014, How Darwinian is cultural evolution? and Dan Sperber's book Explaining culture

Tonal languages and genes: Dediu & Ladd 2007, Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin

The Edinburgh Language Evolution Group: see e.g. Kirby et al. 2008, Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language

Language as a "culturally transmitted replicator": Pagel 2009, Human language as a culturally transmitted replicator

Rate of changes of words is correlated with their frequency of usage: Pagel et al. 2007, Frequency of word-use predicts rates of lexical evolution throughout Indo-European history

"My general perspective…": Acerbi & Mesoudi 2015, If we are all cultural Darwinians what's the fuss about? Clarifying recent disagreements in the field of cultural evolution

Comments

***

Thom Scott-Phillips: What's the fuss about cultural attraction?

20 years ago evolution was a "dirty word" in linguistics (see McMahon, 1994, p. 314). Much has changed: the cultural evolution of languages is now a thriving area of research. There are now many findings about the factors that influence how and why languages change in the ways that they do, and in particular how and why systems of conventional codes begin to acquire the sort of features that identify them as languages. In chapter five of *SOM*, I suggest that these findings are best framed within the context of Cultural Attraction Theory (CAT). In the comments above Alberto continues a productive discussion on this topic, which he has had with myself and others over the past year or so (see previous entries on his blog, and his recent paper on the topic). At the centre of this discussion is the relationship between CAT and other frameworks that have been used to study cultural evolution. Let me in this response sketch one of the reasons why believe that CAT is the right framework for explaining culture. This will allow me to then respond directly to some of Alberto's queries.

For many disciplines, the key intellectual question is why the world is the way it is, and not some other way. In physics, for instance, this leads to the search for physical laws. In linguistics, this question sometimes goes by the name of Greenberg's problem: why are languages the way they are? Biology asks why organismic form is the way it is, and the theory of natural selection is so important exactly because it provides an answer to this question (see below).

By analogy, a theory of culture should aim at explaining why cultural items take the forms that they do. The items in question are mental and public representations, each of which can be cultural to different degrees. So, the goal is to find a way to explain the relative distributions of tokens of different types of cultural representations, both public and mental. This is the agenda of what has been called an epidemiology of representations.

CAT is a proposed answer to the challenge of an epidemiology of representations. As Alberto summarises, the central claim is that in the process of cultural propagation, factors of attraction tendentially cause cultural items to be modified in non-random ways, such that some cultural representations become common in a population, and some do not. If this is correct, it provides a genuinely causal framework with which to explain culture. More specifically, it implies that to explain why a given representation (or set of representations) is common (or not) in a population, we must identify the corresponding factors of attraction.

Let me elaborate on the importance of this point, through a comparison with the theory of natural selection. The idea that there is often a good fit between organism and environment predates Darwin. Darwin did of course contribute a great deal of empirical data and rigour to this finding, but the contribution for which he is most celebrated is that he identified a specific process that links the two sides of this observation. Specifically, he proposed a process by which one side can, in a largely systematic way, cause the other to take certain forms, and not others.

This finding motivates a number of research agendas, of which I here highlight just one: adaptationism. Once Darwin described the causal link between form and environment, he justified explaining organismic form by reference to the corresponding selective environment. This means that for any given trait of interest, the explanatory challenge is to identify the corresponding selective environment(s) in which the trait evolved and is maintained (including the possibility that there are no such environments). Over the past 40 or so years, behavioural ecology, among other disciplines, has achieved vast empirical success pursuing this agenda.

With this background in mind, let me state simply why I believe that CAT is the right framework for language evolution, and indeed for explaining culture in general: it fulfils the same epistemic function for culture that natural selection does for biology. That is to say, it provides a causal, mechanistic explanation of why cultural items should be expected to gravitate in particular directions, and not others. Just as with natural selection, this finding, assuming it is true, motivates a number of research agendas, including one that roughly corresponds to adaptationism. We could, somewhat clumsily, call this agenda "attractionism". The basic idea is that the prevalence and stability (or absence) of some cultural item can be explained by identification of the corresponding factors of attraction.

Language evolution has pursued attractionism with some success, albeit without any reference to CAT. I listed some examples in a preview of SOM that I wrote for the language evolution blog Replicated Typo. However, these findings are not presently unified within one general framework, and it is this problem that CAT can address. Moreover, by framing language evolution in this way, we can, as I said in the précis to this book club, link explanations of structure in language with explanations of structure in other cultural domains.

Let me finish with some quick replies to three specific and important suggestions that Alberto makes. My hope is that the above background will help to make transparent my perspective on each of these.

> "…when domains are close to the 'preservative extreme', it is useful, for pragmatic reasons, to consider them as standard evolutionary systems."

I agree that this may be a useful idealisation in some cases. But we should nevertheless recognise that cases of cultural propagation that are largely or exclusively preservative still lie on a continuum with other cases. Any truly general framework should address the whole continuum, and this is what CAT aims to do.

> "More than asking which aspects are in general more important, it may be more productive to ask when and why transmission is preservative or reconstructive, and what the consequences are for the resulting cultural dynamics."

I agree entirely. This is part of the agenda of attractionism.

> "A further step would be to identify which features of languages are due to cultural attraction forces and which features are due to processes included in 'standard' cultural evolution models…"

This is a false dichotomy. The processes described in "standard" cultural evolution models are a subset of cultural attraction in general.

Mathieu Charbonneau: Cognitive constraints on language

As a man with no horse in this race, I have no problem in accepting both Alberto's replicator view of cultural stability (which should really be relabelled as something in the like of the "conservative" approach, since it is clear since Henrich & Boyd's [2002]) paper that the conservative approach needs not rely on cultural replicators) and the cultural attractor view of cultural stability (I found the Claidière & Sperber [2007; 2010] papers right on point as a clarifying and viable alternative to the conservative view). As long as both approaches are used to study the stability of some tradition, acknowledge one another as a relevant alternative causal-mechanistic explanation, and pay due respect to the real psychological processes underlying the stability of the studied tradition (and do not revel themselves into general mathematical models with little empirical grounding), I see no deep opposition between these views. Which one set of mechanisms is supposed to be the most important – faithful transmission plus discriminatory selection vs. transformation circling around a nearby cognitive gravitational point – appears to me as a problem for a case-by-case empirical investigation of specific traditional stability, but not one for a more general, united theory of cultural evolution. I do not want to partake into this debate, and candidly (and quite frankly), I am still a bit puzzled why there is such a debate in the first place.

But this is not my point. Rather, I would like to suggest a third possible way that traditions might stabilize, and, as I hope to show, whereas it seems to fail to explain *actual* language stability, the suggested "mechanism" might have some consequences for early (proto-)language evolution, both in phylogenetic and ontogenetic processes.

As Thom rightly claims in his response to Alberto's comment: "For many disciplines, the key intellectual question is why the world is the way it is, *and not some other way*. In physics, for instance, this leads to the search for physical laws. In linguistics, this question sometimes goes by the name of Greenberg's problem: why are languages the way they are? Biology asks why organismic form is the way it is, and the theory of natural selection is so important exactly because it provides an answer to this question." (italics added)

I am not sure I agree with how Thom deals with the question, however. It is one thing to explain why things are the way they are given other possible ways they could be, and another to examine the range of ways things can be in contrast to ways it is not possible for them to be. In physics, it is a puzzle just how to explain why the laws of physics landed on exactly the right kind of speci-

fic values they obey to in contrast to other values (as any slightly different values would not have led to any recognizable universe).

This kind of question needs not be set in a high-level, quasi-metaphysical framework as in the case of why the physical laws are calibrated the way they are. Something of a similar puzzle can be found in evolutionary biology. For instance, in the last 30 years or so, an important question in evolutionary developmental biology (evo-devo) has focused not on why one possible, usually observed variant, strove whereas another observed variant failed to be stabilized (natural selection is usually enough to settle these issues). Rather, a different, key question is asked. Simply: "Why were either known forms possible in the first place, and moreover, in contrast to other, non-observed forms?" In a more localized, biological scenario, the question translates as: "Why are some variants of a given trait (e.g., mutations) possible whereas others are not?" As far as I know, neither the conservative view (defended by Alberto, among others) nor the transformative view (defended by Thom, among others) of cultural stability address this question, neither in fact nor in principle. I would like to suggest that the "internal view" (see below) I suggest here offers an alternative explanation to cultural stability to both conservative and transformative mechanisms, and that it is an important part of any process of cultural evolution (language evolution included), as it is the only one equipped to explain why some cultural (language) forms are not possible.

For evo-devo, the question of why observed forms are possible in the first place is a central one. Major advances in evolutionary biology have shown that what might look like natural selection (conservation or attraction in the cultural cases) is merely due to constraints in the sorts of phenotypes that *can be produced at all*, i.e., even before any external stability process can act on these variants (a phenomenon referred to as developmental convergence; see McGhee, 2007; McGhee, 2011). A biological form might be stable enough only because it cannot evolve any more. Natural selection or any cultural analog process have nothing to do with such stability. It is simply the space of possible producible variants that restricts what can be stably produced and reproduced again, and which other forms cannot (see Charbonneau, 2015 for a cultural analog of this process). This means that a biological form might be very stable only because of the set of physical, developmental, and material constraints they evolve in forces them to stop evolving, and thus stagnate, without any adaptive mechanism being required to explain their stability. This latter form of stability is studied in evo-devo as an impediment to the evolvability of species (Altenberg, 1994; Altenberg, 1995; Wagner & Altenberg, 1996). Such "internal" constraints can have major phylogenetic effects. (I say they are internal only because they are not the effect of the immediate selective environment of the trait but merely the effect of the space of *producible* variants (for selection to act, the phenotype needs first to be produced).)

To get back to the evolution of languages, because this is what we are discussing here, I would be very curious to read any participant's thought on what sorts of constraints might limit languages to evolve in a limited space of variation. I am no language expert but, for instance, I have always found that the claim that modern human languages are infinite in scope fundamentally absurd (Thom suggests, with others, that the infinite scope of languages is a qualitative difference, whereas coding systems can only account for quantitative change, and that this serves as a genuine phenotypic difference-maker between what the code model in contrast to the ostensive model can explain, [p. 47]). Of course, assuming some sophisticated generative mechanisms as those of Universal Grammar, you can produce – in principle – an infinite scope of propositions. But constraints on working memory (not to mention any human being's limited life span) will limit the complexity of the sentences anyone can *actually* produce. I dare anyone to demonstrate we can produce an infinite set of sentences, and thus what such a phenotype would actually look like.

I agree that, when considering modern human languages, there might not be any obvious constraints on just what propositions we can produce, but to me this seems more the effect that we cannot imagine a more complex form of language that we are in fact capable of. This observation might seem trivial at first, but I would argue that when this limitation is set in the context of the evolution of languages (for instance what earlier forms of languages might have looked like), it might have some important evolutionary role to play.

Consider the following. Newport (1990) argues that the differences in language acquisition proficiency at different maturational stages might be an effect of the cognitive *constraints* in the learner's working memory. In a nutshell, the older you are, the less proficient you are in learning a language for the first time because your other cognitive capacities are too sophisticated to learn the language in the "normal" way. This is supposed to be a consequence of the fact that ordinary language acquisition depends on a bottleneck-effect imposed by the limitations of the early cognitive capacities that infant of the right age find themselves constrained by. In other words, it is good to have limited cognitive capabilities early on in your development because those constraints simplify what you need to learn, and as your cognitive capabilities increase throughout your development, you can incrementally build a more sophisticated understanding of the language you are nurtured to learn.

When cast into an evolutionary scenario (assuming ontogeny recapitulates phylogeny), this succession of developmental constraints suggests that early humans were only capable of learning a proto-form of language in a very restricted space of language variation. This is because their cognitive abilities could not allow for much linguistic variation in the first place. If there were any cultural evolution at such early stages, "cultural" language evolution might appear very stable if only because there were not so much possible variants to evolve into in the first place. In such populations, we would observe stable proto-languages if only because of the cognitive constraints. However, as cognitive capacities increased throughout our lineage (e.g., as a larger working memory evolved), the range of possible languages and their range of possible complexity might have broadened. What started with a uniform "proto-language" – uniform not because of conservative nor transformative mechanisms, but because of internal constraints in the space of possible language variants – became more and more subject to cultural evolution as the space of possible language variants broadened.

Agreed, the "internal" constraints I have pointed at here might not apply to modern languages. Again, I am no language expert. I am only suggesting a different way to look at cultural stability. I have argued elsewhere that such constraints might explain stability in other areas of cultural evolution, especially technological evolution (Charbonneau, 2015). Nevertheless, I am very curious to know if any such constraints might apply to modern language capabilities, as a cause of cultural stability that neither falls under the conservative paradigm suggested by Alberto nor the transformative scenario of Thom.

6

NO COMMUNICATION WITHOUT REPUTATION, NO REPUTATION WITHOUT COMMUNICATION

By Olivier Morin

I chose to organise an ICCI book club around *Speaking Our Minds* because it is an exceptional book in more than one way. It ties together two research traditions – the pragmatic approach to linguistics and the Darwinian legacy in biology –, that lie at the heart of our field. It does so in a perfect format – the book is a delight to read, and to teach. Yet a good Book Club book should be more than a good book: it should also stir up discussions.

Speaking Our Minds is sure to do that: here is a book that claims to solve every major problem linked with the birth of human language! I won't be the only one who doubts that. I think that Thom Scott-Phillips' account of how human communication evolved leaves at least one key problem unsolved: How is communication stable in the face of free-riding and deception? In the two sections of *SOM* that he devotes to that issue (6.5–6.6), and in his paper on the correct application of animal signalling theory to human communication (Scott-Phillips, 2008), Thom rejects the standard solutions to the problem. In the field of language evolution, these tend to come from signalling theory, and to involve the Handicap Principle. With some nuances, I share his grim view of the prospects of signalling theory. Yet his own solution is not just extremely sketchy, at times it verges on circularity

Words are cheap: The Handicap Principle is a red herring

Let's start with one thing we all agree on: one family of explanations that work very well for biological signals in general cannot explain the stability of human communication. Here I dwell on Thom's opposition to Zahavi's much misused "Handicap Principle". At the principle's core is the idea that a signal's honesty may be stabilized by the costs involved in displaying it. The peacock's tail, for instance, is a wasteful ornament, so wasteful that the fact of growing it is an indicator of good health. Another popular example (indeed, handicap theory's poster child) is the springbok's pronking behaviour. Springboks, it is said, will jump and frolic in front of a predator simply as a way of advertising their athletic

capacities: what springbok would be mad enough to nag a lion, unless it is extremely fast?

Thom cites the peacock's tail, but not the springbok example. I think that is not fortuitous: pronking springboks do not really obey the Handicap Principle as Thom sees it. In his view, the Handicap Principle involves signals that are costly to produce, and that is not the case of pronking. What makes pronking a handicap is its consequence: the risks the animal incurs by frolicking in front of a predator, wasting time and energy that it might need to run for its life. The cost of taking those risks is not paid when producing the pronking signal. It may be paid later when the lioness chases after the springbok. The costs of pronking, in other words, are deterrents: they are penalties that a springbok would incur if it were signalling dishonestly. They lie in the future. In the peacock's case, in contrast, the costs are (in large part at least) paid by the peacock when it grows its tail, as a necessary aspect of producing it. In Thom's view, it seems that a genuine handicap must entail such production costs, and so he reserves the term "handicap" for these. This clarification is one of the book's most welcome contributions (more on this below).

Conspicuous consumption and the Handicap Principle

Thom (§6.5) has two points to make against handicap-principle explanations for the evolution of biological signals, when humans are concerned. First, he blames much of the literature for failing to understand that the Handicap Principle requires differential costs: a handicap must be "[f]ree to those who can afford it, very expensive to those who can't" (p. 146). In that vein, Thom criticizes those who construe conspicuous consumption as an instantiation of the Handicap Principle, because conspicuous consumption implies costly purchases, but no differential cost:

> "…if you can afford to spend money on objects like diamond rings and designer clothes, whose non-signalling functions could be equally well fulfilled by far less expensive purchases, you must be wealthy indeed.…However, [this] is not a great example [of the Handicap Principle] because it is not, in fact, a handicap at all. This is because there are no differential costs involved. Put simply, jewellers do not have different prices for different customers…Conspicuous consumption is an index of wealth, not a handicap." (p. 148)

I beg to differ. Marginal utility theory teaches us that spending ten thousand dollars on a watch is indeed more expensive for us than it is for Bill Gates. The dollar that gets my fortune to grow to one hundred and one dollars is worth more than the one that makes it grow from one million to one million and one. The costs of a one dollar purchase are correspondingly lower for the very rich. Thus, conspicuous consumption is a true handicap: an expensive watch is a purchase that the rich can afford at a lesser opportunity cost than the poor, and it is, of course, causally associated with wealth.

Similar remarks could be made about some other examples mentioned in the book – deer calls, for instance. A deer call's pitch (or its formants distribution, but I'll keep things simple) reflects the caller's size: having a larynx of a given size produces a call with a certain pitch. Using Maynard-Smith and Harper's terminology (Maynard-Smith & Harper, 2003), Thom calls this an index, that is to say, a signal that is reliable because it is physiologically impossible to fake it. In Thom's view, indexes cannot obey the Handicap Principle, because they involve no differential cost (unlike, say, the peacock's tail). Things are not so obvious to me. First, deep-pitch calls do have a physiological cost (the cost of growing a large body and giving out calls). Second, it is not completely unimaginable that a small deer might grow a big larynx, with terrible consequences for the deer's viability and fitness. So, rather than being impossible, the low-pitched call is immensely costly for small deers, virtually free for bigger ones. "Free to those who can afford it, very expensive to those who can't": the essence of differential cost. In this deer calls are no different from peacocks' tails.

(At this point, one might reply by citing Maynard-Smith and Harper, for whom a handicap must always include a "strategic cost" in addition to its "efficiency cost". In other words, a cost must be incurred over and above the cost of actually transmitting the relevant piece of information. Yet I think there are reasons to resist this view. First, because it makes little sense of cases like the peacock's tail: in a way the peacock pays exactly the cost that needs to be paid to advertise its own health – it is not being inefficient. Second, strategic costs jar with the "differential cost" principle – the view that handicaps should be "[f]ree to those who can afford it".)

Knowing a handicap from a deterrent

On the other hand, I fully endorse Thom's second point against signalling theory. *SOM* (p. 147–150) does a fantastic job of pointing out the muddles that people fall into when they fail to distinguish handicap costs from deterrents. As Thom points out, handicap costs are "paid by honest signallers…as part of the very act of being honest", while deterrents are "paid by dishonest signallers when they deviate from honesty" (p. 148). The distinction matters when discussing evolutionary stability, because handicap-based signals are likely to prove more reliable than the ones whose costs derive from deterrence. The peacock's tail indicates health simply by virtue of its being there. The springbok's pronking, in contrast, proves little by itself, when performed in an empty savannah. It is a reliable signal of health only if (i) a predator is present, (ii) that predator is fast and hungry, (iii) the springbok knows this and (iv) the springbok is correct in its prediction that the lion can easily be outrun. If any of these conditions fails to obtain, the deterrents collapse, and so does honesty.

(Thom's discussion brought back quite a few memories of discussions with handicap aficionados who failed to see this crucial difference. In one such discussion (available online here, here, and here), a group of behavioural ecologists argued that almost everything about diplomacy and international relations is based on the peacock's tail principle. They refused to see the difference between a

nuclear test, which is costly to produce and hence a reliable indicator of military capacity, and a declaration made, say, in front of the United Nations, which might involve unpleasant consequences if one lies, but is nothing costly to utter and thus proves nothing by itself.)

To summarise, I have minor disagreements over two points of interpretation of the Handicap Principle: I think that the Handicap Principle might originally have meant to encompass deterrents (as the popularity of the pronking springbok suggests). I also think that differential costs are nothing special, and because of this I wouldn't restrict the scope of the Handicap Principle as narrowly as Thom does: Veblen's conspicuous consumption is a bona fide illustration of the Handicap Principle, and there is nothing deeply wrong in linking natural indexes (like the pitch of deer calls) to handicaps.

More importantly, though, I fully agree with Thom's main point: the Handicap Principle cannot help us understand the stability of human communication, because its proponents make a muddle of the distinction between handicaps and deterrents. The former kind of cost plays no role at all in making language reliable. Words are cheap—meaning that the cost of uttering them is (i) trivially low and (ii) unrelated to what the words signal. Handicap costs are not what makes human communication reliable. What does, then, must be some kind of deterrent.

Thom's solution: the reputational costs of deception

Having started on such an excellent footing, Thom's account essentially stops where it should start. Human communication is kept stable and reliable, in his view, because of the loss of reputation that deception would entail. On this, Thom's 2008 paper is slightly less elliptic than the book:

> "Sufficient conditions for cost-free signalling in which reliability is ensured through deterrents are that signals be verified with relative ease (if they are not verifiable then individuals will not know who is and who is not worthy of future attention) and that costs be incurred when unreliable signalling is revealed. These conditions are fulfilled in the human case: individuals are able to remember the past behaviour of others in sufficient detail to make informed judgements about whether or not to engage in future interactions; and refusal to engage in such interactions produces costs for the excluded individual…Moreover, this process would snowball once off the ground, as individuals would be able to exchange information – gossip – about whether others were reliable communication partners (Enquist & Leimar 1993); and that exchange would itself be kept reliable by the very same mechanisms." (Scott-Phillips 2008)

This view, I think, takes too much for granted. We would need to know how humans came to have mutually beneficial interactions, on a regular basis, with individuals who could cheat them, but most of the time did not; how humans could reliably verify the validity of others' signals; how, when a signal seemed inaccurate, one could know that one was the victim of a deceptive informant (as opposed to

an incompetent one). Of course, a proper answer to all these questions cannot involve communication, at least not in the first few steps; nor can it appeal to the kind of cooperation that communication makes possible.

From the few lines above it seems that Thom has in mind a scenario involving mechanisms of partner choice. Redouan Bshary's work on mutualistic cooperation among cleaner fishes (see for example Bshary & Grutter, 2005) could serve as an analogy. These fishes need to coordinate on mutually profitable cleaning operations (removing dead skin, parasites, and such like). They appear to memorise it when one fish has assisted them longer and more efficiently than others in the past: they preferentially choose that fish for future joint ventures. Fishes who fail to cooperate optimally, for instance by giving up cleaning before they are expected to, are gradually pushed away from the pool of desired partners. One thing to say for this scenario is that it does not involve communication in any serious sense: memory of individual encounters, occasionally augmented by eavesdropping, does everything.

Yet our ancestors were not cleaner fishes. I mean that, before the rise of communication, nothing guarantees that they routinely recruited non-relatives for mutually profitable ventures. We cannot be sure, either, that failures to cooperate were visible to the public eye like they are in cleaner fishes. There is a bigger story to be told here, but it is not told in *SOM*. Suppose, however, that we skip this issue. Suppose that mutually beneficial cooperation and partner choice were well established before our story starts; could they explain the early stability of communication?

I am not sure. For a number of reasons, liars should be much harder to spot than cleaner fishes who free-ride on others' efforts. Useful communication is about things that are far from us, in space or time. Because of this, communication's accuracy can seldom be judged here and now, in front of numerous witnesses, like the failings of cleaner fishes. The consequences of a lie for its victims are typically delayed, and difficult to interpret. Suppose I told you that yonder tree is full of juicy apples, but you find out (moments later) that it only carries poisonous manchineel fruits. Can you call me a liar? Not before you make sure that (i) I was being malevolent, not just incompetent, (ii) you yourself are competent enough to tell an apple tree from a manchineel. Even if you succeed at steps (i) and (ii), the fact that I lied to you is not (unlike cheating in cleaner fishes) a public event that people can eavesdrop on. Thus, on the face of it, we should doubt the power of partner choice, individual memory, and eavesdropping, to prevent deception.

A more reliable alternative would be public reputation, established through communication between individuals. Thom clearly has it in mind in several parts of the book (e.g., when he cites *Othello*'s Cassio); but the circularity of this solution is too obvious to stress. Public reputation is kept reliable by the very acts of communication that it is supposed to police. The messages that make up public reputation can, after all, be just as deceptive as any other kind of communication. They can arguably be even more so, since reputation is often diffuse, informal, and not assignable to one source in particular. If we solve the problem of reliable communication by assuming reliable public reputation, I doubt the buck can ever stop.

Until such questions are addressed, reputation will remain the deus ex machina that intervenes at the end of the play, to tie up the loose ends of the plot. What a good plot, though.

Acknowledgement: I thank Denis Tatone for his precious input on this text.

Comments

***

Hugo Mercier: Springbok pronking and early liars

I was under the impression that pronking referred to a rather more specific behaviour than what you described, namely, simply jumping straight up as high as possible. Such a behaviour would then seem to be an index: only a gazelle that has the strength and stamina to jump that high can actually do it. In this case, deterrence would have nothing to do with it. (Note that even if this were the case, your remarks about the deers would apply: gazelles could presumably evolve that are good at jumping high but crap at running, and it's only the cost of this option, and not its supposed sheer physical impossibility, that would keep things in check). Is there any evidence either way?

Another detail: I'm not quite persuaded by your example regarding the supposed difficulty of using reputation as a deterrent to keep communication stable in the early stages of language evolution. It seems that neither (i) nor (ii) are necessary. Regarding (i), I don't see why it would be necessary to establish malevolence. If someone has given you information that turns out to be harmful, you trust them less, that's the end of it. That's enough to create an incentive for them to be more careful when communicating. Regarding (ii), even if you were able to tell the difference, wouldn't you still have a reason to lower the trust you put in the other guy?

And I don't see why a lot of communication wouldn't take place in at least semi-public spaces (i.e., the communication would be aimed at one individual, but others would be likely to overhear), and even why this would be necessary: if an individual often provides harmful information, no one will trust him, and his fitness will be substantially reduced, no?

Olivier Morin: Reply to Hugo

About pronking

Hugo's interpretation is one among a dozen conflicting interpretations of pronking/stotting, and it conflicts with the handicap-principle explanation. Obviously I don't claim to know which interpretation is true, it was just an illustration (I can't tell a springbok from a Thompson gazelle, and everything I know about those I learnt from David Attenborough!).

About spotting early lies

I raised four problems for Thom's account. It has to explain (1) how mutualism arose, (2) how people tell deceiving informants from incompetent ones, (3) how people know that their informants are more knowledgeable than they themselves are, and (4) how knowledge of deception spreads without communication when opportunities for eavesdropping are limited.

I agree that (2) is not the worst problem. People might just avoid useless informants, be they incompetent or malevolent, without trying to know which is which. Yet this is a very imperfect solution. Deceptiveness is easily thwarted by mechanisms of partner

choice or partner control, but it is hard to see how the same mechanisms would "cure" cluelessness. Uninformed people remain uninformed no matter what: incentives just cannot turn clueless informants into good ones. Responding to clueless informants by avoiding them in the future will be a poor move in many cases (next time, the same clueless informant might possess a precious piece of information, and it won't pay to shun her), whereas shunning deceivers will be the thing to do. In a partner–choice framework, it makes adaptive sense to distinguish mistakes from lies. (Besides, Thom's account is explicitly framed as an account of how communication roots out deceptive misuses – "the boy who cried wolf".)

Concerning (3), I don't really understand your objection. If I am facing what I think is a manchineel tree, and the informant tells me it is an apple tree, the only way I can truly know who got it right is by eating the fruit – in which case, if I am right, I am dead.

Anyway, that still leaves issues (1) and (4).

Hugo Mercier: Incompetence and deception

I think that actually most (or at least a lot) of what people have to be careful of in communication is neither outright incompetence nor outright deception, but rather negligence. For instance if the informant in the story says that X tree is an apple tree, she might just be saying that without being really sure (it is something she heard, say), and with no intention of deceiving you. She just doesn't care much whether it's true or not. That might not be a sign of incompetence from her point of view, if she has no intention of going to that tree anyway. But if she gives you this information, she's being negligent in a way (especially if she fails to specify that she is not sure at all). In such a case, it makes sense to 'punish' the informant by trusting her less: she might not learn how to get better information (i.e., become more competent), but she will learn to be more careful when transmitting it. If it is true that such situations arise pretty often, then 'punishment' would be quite a powerful way to insure the stability of communication.

Thom Scott-Phillips: Two unresolved issues around stability

I am of course delighted that Olivier chose to organise a book club around *SOM*, and I would like to thank him here for the work he has put in to make this happen. We are just a few days into it, and it is already abundantly clear to me that this format is sure to advance discussion in several important ways. Olivier's own contribution is a perfect example of this. These advances would not be occurring if not for Olivier's efforts.

Olivier summarises very well my discussion of the evolutionary stability of human communication. My main objective in these sections was to stress that evolutionary stability is explained in a really rather simple way, once the underlying theory is properly understood (it has too often not been properly understood in the past). In short: human communication, or at least the overwhelming majority of it, is kept stable by reputation.

Olivier's comments raise, first, a question about a terminological detail of this claim, and second, a collection of important tangential questions about how such a system could have emerged in the first place. Let me comment, more briefly than I would like, on these two issues in turn.

I argue in *SOM* that conspicuous consumption is not a handicap, since it does not involve differential costs. Olivier disputes this. The costs of conspicuous consumption are not differential in absolute terms, but they are, he points out, differential in marginal terms. The question, then, is whether the costs involved in signal production should be treated in absolute or marginal terms. The answer to this question will come from the modelling literature. Alan Grafen's original model of the Handicap Principle (Grafen, 1990), which lent a great deal of credibility to Amotz Zahavi’s propoals, certainly did conclude that marginal costs could stabilise costly communication. The literature has however moved on considerably since then. Olivier's arguments have therefore encouraged me to revisit some key papers in this area, with this issue, and Olivier's related comments about deer calls, in mind. Doing this, I see that the matter is not straightforward, and needs further consideration. I can right now only promise to report back in due course.

Olivier's questions about the emergence of stable human communication are nicely summarised in his response to Hugo, above: "I raised four problems for Thom's account. It has to explain (1) how mutualism arose, (2) how people tell deceiving informants from incompetent ones, (3) how people know that their informants are more knowledgeable than they themselves are, and (4) how knowledge of deception spreads without communication when opportunities for eavesdropping are limited". This are all good questions, to which (again) I do not have good answers at present. In this case, I think these questions amount almost to a manifesto for an original modelling project, aimed at identifying the range of parameters under which a mass communication system kept stable by reputation could emerge. Done well and comprehensively, a project of this sort should be able to address Hugo's questions about the relative importance of the four problems. It should also tell us if Olivier's suggestion, that public reputation could play an important role, is correct. Certainly, this seems as plausible to me as any speculation I can offer right now. I am slightly saddened that I have little to add to it at present, but I think this is for the best: for me to make additional speculations of any substance would, I think, be little more than second-guessing the results of the project outlined above.

7

COMBINATORIALITY AND CODES

By Liz Irvine

I read this book as part of an interdisciplinary reading group at Cardiff. As we found, there is a lot to agree with in the book, but the commentary below focuses on two points that we found confusing.

Combinatorial communication

Chapter two claims to be about the impressive expressive power of language: we can construct an infinite number of sentences, expressing new ideas and capturing a huge range of meanings with a finite set of building blocks. At least, this is what Pinker and Fodor and others find extraordinary about language. The apparent target explanation of this chapter is whether code communication or ostensive communication is the most likely route to it. But instead of trying to explain the productivity and expressivity of language by focusing on compositionality and systematicity, the chapter focuses on 'combinatorial' communication of a particular kind. However, this neither fits with usual definitions of either combinatoriality (combining meaningless units) or compositionality (combining meaningful units where the meaning of the whole is composed of sub-meanings of parts), and I think ends up answering a very different sort of question.

So, combinatorial communication as defined in the chapter is where two (or more?) meaningful signals are combined to form a new signal whose meaning is not the sum of the meanings of its subparts (p. 27). Instead, a new signal is formed with a totally different meaning. So, adding monkey 'pyows' ('leopard!') to 'hacks' ('eagle!') results in a 'pyow-hack', which means that the group will soon move to a new location. This is not therefore a case of compositionality or combinatoriality. Scott-Phillips admits that this is not really a case of 'combination' either (p. 50), since this is effectively just adding two holistic signals to form another distinct holistic signal, rather than a 'combinatorial' signal.

As far as I can tell, the most plausible way to interpret the claim in this chapter is that it is difficult to combine signals to (non-compositionally) form new signals, if there is not something obvious in the environment which correlates with the meaning of the new signal (so for which the new signal cannot

be either a cue or a coercive behaviour). This means that building up new, non-compositional, vocabulary by adding together existing signals is unlikely under code communication.

This seems plausible – but it is hard to identify just what question this chapter is actually addressing. The massive flexibility and expressivity of language has very little to do with whether we can add existing signals together and get a signal with a totally different meaning. Instead, it is usually taken to be related to our ability to add signals together to form new signals whose meanings are composed of the meanings of its sub-parts, and how we can add those together further in systematic ways to form e.g., sentences where again, the meaning of the whole is at least strongly related to the meanings of its sub-parts. What Scott-Phillips instead seems to be addressing in this chapter is a form of vocabulary building – how to get new holistic signals from existing holistic signals, whose meanings are all unrelated.

The section on ostensive communication also seems aimed at vocabulary building – if you are good at ostension, then you can come up with signals for whatever meaning you want. Here though, instances of "combining" (e.g., p. 43) really are instances of compositionality – new signals are generated whose meaning is composed of the meaning of its sub-parts. Here then, (a) we are not always comparing like with like, and (b) the role of both codes and ostension are related specifically to vocabulary building, not directly to the compositional nature of language.

So, while vocabulary building probably is easier under ostensive communication rather than the code model, and so ostensive systems might be more expressive in this sense, the compositionality of language and so its impressive productivity is not addressed. Further, to the extent that proto-language demands compositionality (does it?), this then does not show that code communication is inadequate to build proto-language. .

The relationship between ostensive-inferential communication and communication via codes

This is obviously supposed to be the theme throughout the book: that you cannot get to language via an elaboration of the code model, even by adding in ostension. Instead, the claim is that in some sense, ostensive communication is primary, and made more expressive via conventional codes (e.g., p. 16, and elsewhere). However, I found it hard to track what exactly this claim amounts to throughout the book. (Is it a claim about the actual development of linguistic systems? A conceptual claim? Just a plea for a shift in research attitudes?)

The best I reconstruct the claim is as this: code communication with ostension tacked on is not as flexible as language actually is; to get real flexibility, you have to start with ostension first, which is helped along by conventional codes. Given the claims made throughout the book that language is not code-like anyway (using ostension you can get words to mean whatever you want), this seems like a straightforward statement. But I do wonder how this plays out in the actual development of early linguistic systems. There it seems less straightforward.

First, the 'code-like' features of language are incredibly useful. Even the metaphorical or more flexible uses we put language to are often based on nets of semantic associations. Further, while we can in theory use words in radically flexible ways, a huge amount of communication does rely on meanings being fairly stable (conventional codes). This is because codes not only make communication more powerful, they also make it (cognitively) much easier.

This seems particularly relevant in the development of early linguistic systems, where, given that the first forays into ostensive communication were likely to be hit and miss, you'd need all the help you could get. Communicating with someone in the absence of a shared language is hard work, even in fairly simple contexts, and even if you both have A+ mind-reading abilities. What you need are ways of minimizing ambiguity to a level where mind-reading has a reasonable shot.

Accordingly, one way of minimizing ambiguity in communication that is often discussed in language evolution is iconicity (e.g., here in chapter five). Iconic signals 'look like' the things they represent, which should make it easier to grasp their meaning. Discussions of the important role that iconicity may have played in early communication systems is predicated on precisely the idea that linguistic communication is hard, even with ostension, so signs probably started off iconic and then later became arbitrary.

Another way of minimizing ambiguity is of course with codes. It then seems reasonable that at least some conventional codes were likely to have been derived or adapted from existing natural codes, with others added on via different means (e.g., perhaps conventionalized iconic signs). In this case, there would have been some 'continuity' (p. 48) between earlier code communication and later ostensive communication. Indeed, one of the questions that kept coming up in the reading group was what happened to earlier natural codes – surely hominids would not just drop them entirely, but, as great apes do, use them in ever more flexible ways. Something like this reliance on codes is also found in e.g., §5.4 – where proto-language includes a set of "more-or-less stable communicative conventions" (p. 117).

In this case, early language users would presumably have relied on, among others, both complex coding/decoding mechanisms (association) and mind-reading abilities (metapsychology) to get early proto-linguistic systems off the ground – not just one of them in isolation. On a conceptual level, and rather obviously, all you need for ostensive communication is ostension, but practically, associative mechanisms would also have been crucial in getting linguistic ostensive systems going.

However, it is hard to tell if this picture where both mechanisms have a role to play, and neither route (pure code or pure ostension) in isolation looks particularly plausible, amounts to a challenge to Scott-Phillips' view. Perhaps the difficulty in identifying claims here (and elsewhere in the literature) is based on ambiguities in the explanatory roles that particular mechanisms are supposed to play, and in exactly which stage of the evolution of language. Clearly, if the primary marker of the difference between linguistic communication and non-linguistic communication is deemed to be the use of os-

tensive-inferential abilities, then these will play a key role in explaining the emergence of "language proper", but if linguistic communication is identified in some other way (e.g., displaced reference), then the focus may well be elsewhere.

Comments

***

Thom Scott-Phillips: How is ostension prior to the linguistic code?

I am of course delighted to hear that Liz and her colleagues discussed *SOM* at length in their reading group, and I am pleased to have this opportunity to address the concerns she raises.

In her comments on combinatorial communication, Liz suggests that my attention is not focused on exactly the right topic. She points out linguists often stress the importance of two particular types of combination: combinatoriality (combining meaningless units), and compositionality (combining meaningful units into a whole, whose own meaning is a composition of the meanings of the component parts). I focus instead on what I call combinatorial communication: combining meaningful units into a whole whose meaning is not simply the sum of the meanings of the component parts. Liz believes that because I focus on this, rather than on combinatoriality or compositionality, it is at least somewhat unclear what my discussion tells us: "it is hard to identify just what question this chapter is actually addressing".

Let me therefore try to clarify. One of the main goals of chapter two was, in a sense, a negative one: to argue that a communication system that has anything like the sort of combinatorial richness that languages do is extremely unlikely to emerge if it is based upon natural codes. To make this argument I focused on what I call combinatorial communication. I did this for a variety of reasons, one of which is that combinatorial communication is, more-or-less by definition, the most simple way in which meaningful elements can be combined. My question was: how far can you get combining natural codes together? And my answer is: not very far at all. Certainly nowhere near anything that looks like a language. If I am right about this, then the fact that combinatorial communication is not a typical category for linguistic analysis is essentially beside the point. You just can't get there from here. You have to look elsewhere for a way to get a widely combinatorial system of any sort off the ground.

Let me put the point another way. Liz writes that "The massive flexibility and expressivity of language…is usually taken to be related to our ability to add signals together to form new signals whose meanings are composed of the meanings of its sub-parts, and how we can add those together further in systematic ways to form e.g., sentences". This is true, but it misses the point. The combinatorial power of natural languages does increase expressivity, but this expressivity is parasitic on ostension and inference. Many animals can, I assume, combine things together. There is no reason to think that it is computationally difficult. (This is, incidentally, why I explored experimentally the possibility that bacteria use combinatorial communication.) So combining things is nothing special in and of itself. It is only special when it sits atop ostension and inference. The usual view that the expressivity of language is due to combinatorics is only partly true, and the part it misses is, I insist, critical.

This point leads directly to Liz's question about the sense in which I claim that ostensive communication is prior to the linguistic code. She asks whether this is "a claim about the actual development of linguistic systems? A conceptual claim? Just a plea for a shift in research attitudes?". All three, is the answer. It is first a conceptual claim: the linguistic code just is a set of conventional codes built precisely in order to increase the expressivity of ostensive communication. But what I was also concerned to argue throughout *SOM* is that this conceptual point translates into, yes, a claim about the evolutionary development of languages: ostension must come first. And for both these reasons, and others, I do think that some shift in research priorities would be healthy for the language evolution community.

All these issues come together in Liz's assertion that "another way of minimizing ambiguity is of course with codes". Much of chapter one of *SOM* was dedicated to making the case that, when it comes to conventional codes – which is what languages are – this is not true. It is common in linguistics to believe otherwise, but a (the?) key lesson from pragmatics is that all linguistic utterances are ambiguous. As I put it in *SOM*, "*as codes*, languages are very defective indeed. In fact, they are wholly ineffectual" (p. 17, italics in original). Put simply, the linguistic code is parasitic on the ability of ostension and inference to handle the inherent ambiguity of conventional codes. This is the reason why I claim that it is prior, in essentially all the ways that matter from an evolutionary point of view.

Liz Irvine: On 'Continuity'

I want to have another go at expressing some ideas that perhaps were not clear in my first post, on the relation between Thom's conceptual and evolutionary claims about language, and how they link up to research attitudes.

So, the conceptual claim (following Grice, Sperber, and others) was:

(1) Linguistic communication is not code-like, but relies on ostensive-inferential capacities of the communicating individuals.

The evolutionary claim, which Thom seems to suggest follows directly from this (or not?), goes something like this:

(2) The evolution of linguistic communication directly tracks the evolution of ostensive-inferential capacities. No other species have these capacities. There is therefore a straightforward discontinuity between the communication systems of non-human species, and the linguistic communication of humans, because no other communication system is 'language-like'. (e.g., "…nothing that looks even remotely like language can emerge prior to the evolution of ostensive-inferential communication." [p. 46]).

So, Thom makes an inference from the conceptual claim about a distinctive feature of language, to an evolutionary claim of discontinuity. My question is what the claim of discontinuity actually amounts to, and whether it is warranted. The worry tracks Bar-On's distinction between synchronic continuity/discontinuity (e.g., what is distinctive about language, compared to non-linguistic communication) and diachronic continuity/discontinuity (whether there is a clean evolutionary divide between non-linguistic and linguistic species) (Bar-On, 2013, p. 342–343). As she points out, synchronic discontinuity is compatible with diachronic continuity, and this is what I want to press on here.

So, to repeat, the inference above moves from a claim about what is distinctive about linguistic communication, to the claim that linguistic communication is evolutionary discontinuous with other forms of communication (nothing else is language-like). But this requires a few other moves.

First, on a conceptual level, one needs to decide what counts as being language-like in an evolutionary setting. But this does not follow automatically from a conceptual definition of language, or identifying its core features. One can agree that ostensive-inferential abilities are essential and unique to linguistic communication, but still think that there are evolutionary precursors that do not rely on these abilities that are 'language-like' in an interesting sense. Asserting that only full-blown ostensive-inferential communication is like full-blown ostensive-inferential communication is fine in a conceptual sense, but is a peculiar assertion to make in an evolutionary setting.

Second, on an empirical level, one then also needs to assess which (aspects) of non-human animal communication systems are language-like. This is an ongoing project.

So, in order to sustain your inference, you need to assume/argue at least the following:

(1) Ostensive-inferential communication requires fourth order metarepresentational capacities, and without these, no communication system is 'language-like'. But Moore has argued that Gricean communication may not require this degree of metapsychology, in which case there are simpler precursors to full-blown Gricean communication. Arguably, these would be 'language-like' in an interesting evolutionary sense.

(2) There are no 'interesting' stages of proto-Gricean communication, and there are no other relevant features of language that might make a communication system 'language-like'. This is not sufficiently defended. For example, Bar-On's account of expressive communication as a relevant precursor to language is dismissed in one paragraph in the book (bottom of p. 46–47); apart from anything else, this is arguably too quick. Further, the worry that there is no 'detailed account' of something midway between code communication and full-blown ostensive-inferential communication is not sufficient to disregard the possibility that there can be one. The reason why detailed alternative accounts are not in abundance is presumably because of the ongoing popularity of e.g., Tomasello's cooperative communication model and the focus on social cognition in general. Yet there is increasing evidence that ape social cognition is not a million miles away from human social cognition (controlling for socio-communicative environment), which in itself suggests that there is some interesting notion of continuity to work with here (see Slocombe's [e.g., Slocombe & Zuberbühler, 2005] and Lyn's [e.g., Lyn, Russell, & Hopkins, 2010] work, Call's review [2011], etc.). Also, the idea that (full-blown) ostensive-inferential abilities are the only important or interesting features of linguistic communication with which to make judgments on evolutionary continuity is perhaps a bit odd. Things like displaced reference, flexible use of signs, shared attention, learning mechanisms, etc., seem relevant too.

(3) The only relevant sense of 'continuity' is in terms of the means or mechanism of communication (e.g., via ostension and inference). But there is another sense of continuity that applies to the linguistic/communication system itself. So, on the assumption that early language use was hit and miss, early proto-linguistic systems would presumably have still used existing codes in flexible ways, presumably made a lot of use of pointing, and on top of this would also have made slow and laborious use of the 'direct route' to generating new signs with fairly fixed meaning. So, in terms of what was actually 'in the lexicon', there may have been at least some continuity in the early stages. This is the sense of continuity that came up repeatedly in the reading group at Cardiff, but is quite different to the one meant (I think). (This is also similar to Bart de Boer's comment).

Now, in one way of course, whether we say that the evolutionary trajectory of linguistic communication was continuous or discontinuous doesn't really matter, as these terms are pretty vague anyway (and with a different emphasis, the same data can support both claims). But when it does matter is when it comes to research programs and attitudes. If, on the one hand, a strong claim of discontinuity is made, then this makes it easier to ignore features of communication systems that might actually turn out to more 'language-like' than previously thought, or to keep to a narrow conception of what language is and what its important features are. Of course the flip side is that if continuity is claimed, then we might dismiss genuinely unique features of linguistic communication.

So, while trying to shift research attention to the core importance of pragmatics and social cognition in understanding language evolution is a good thing (though, following Bart de Boer's comment, this is already a major research area in non-Chomskyan studies of language evolution), it seems like this can be done just as well by emphasising: 'PRAGMATICS IS REALLY IMPORTANT!', rather than making additional and perhaps problematic claims about evolutionary discontinuity.

Thom Scott-Phillips: Reply to Liz

You write, correctly, that "One can agree that ostensive-inferential abilities are essential and unique to linguistic communication, but still think that there are evolutionary precursors that do not rely on these abilities that are 'language-like' in an interesting sense…on an empirical level, one…needs to assess which (aspects) of non-human animal communication systems are language-like".

I entirely endorse looking at non-human communication systems for precursors of full-blown ostensive-inferential communication. In *SOM* I set out tests that would unambiguously identify a communication system as ostensive-inferential; but I did not, it is true, discuss what potential "mid-points" might look like, and how we could identify them. I nevertheless agree, without reservation, that this is an important conversation to have. Mathieu's and Katja's comments are good starting points. Having said that, I would like to stress that what we should be looking for here are behaviours that may be partially ostensive (i.e., are produced with an intention to make it manifest to an audience that one has an intention to communicate). Yet there is almost no comparative work specifically focused on ostension. Yes, there is work on intentionality (shared or otherwise), and this is relevant, but there is very little on ostension, properly understood. I would be delighted to see more work of this sort. (I do know that there is some work on ostension in non-human species, but this is a small subset of research focused on more basic aspects of social cognition, and it is dwarfed by research programs focused on the codes used in non-human primate and other communication systems. This is what I mean by "very little".)

Let me turn now to the three points that you argue I need to assume and/or argue for, in order to sustain my arguments:

> "(1) Ostensive-inferential communication requires fourth order metarepresentational capacities…but Moore has argued that Gricean communication may not require this degree of metapsychology, in which case there are simpler precursors to full-blown Gricean communication…"

As you know, I argued in *SOM* that ostensive-inferential communication is a richly metapsychological activity. You are right that Richard, among others, has made counter-arguments. But I could not rebut those arguments in *SOM* because the detailed version of them was not published at the time. In fact, the detailed version is still not. It is true that parts of the argument exist in some of Richard's papers, but the most explicit versions are still under review at the moment. So how could I rebut them?

> "(2) There are no 'interesting' stages of proto-Gricean communication, and there are no other relevant features of language that might make a communication system 'language-like'…"

As I said above, I would happily welcome proposals of this sort. That is: I would welcome substantive proposals of what partially ostensive communication would look like. There is, again, very little work of this sort. You are right that Dorit Bar-On's work is one exception, but as I said in *SOM*, I do not find it persuasive.

It may also be relevant to point out that I don't think that an author has an obligation to address all arguments contrary to their own view. Richard made a similar complaint to yours, and I stand by what I said in response: I explicitly stated in *SOM* that the agenda was not to provide a comprehensive overview, but to make a positive case for my own arguments. (Perhaps this is a disciplinary difference in expectations: I cannot help but notice that, of the various book club commentators, it is yourself and Richard – both originally trained as philosophers – who have complained that *SOM* does not discuss other positions in enough detail.)

> "(3)…in terms of what was actually 'in the lexicon', there may have been at least some continuity in the early stages."

I agree. But this is not the sort of continuity I argued against in *SOM*.

One last point. You write that "while trying to shift research attention to the core importance of pragmatics and social cognition in understanding language evolution is a good thing (though, following Bart de Boer's comment, this is already a major research area in non-Chomskyan studies of language evolution)". I can't agree. If this is a major research area, why does "pragmatics" barely feature in the indexes of edited collections on language evolution? There are widespread acknowledgements that pragmatics is important, but this is mostly lip-service. Consider, for instance, research on the codes used in non-human primate communication. This is much bigger area than comparative pragmatics and, critically, this research rarely if ever considers whether the codes that are being studied are natural codes or conventional codes. This would not be the case if pragmatics was taken as seriously as it should be.

8

ALIGNMENTS ACROSS DISCIPLINES

By Ira Noveck and Tiffany Morisseau

This book left a very positive impression on us both. It is practically a manifesto for clear thinking about doing proper Gricean analyses in applied areas of communication. *Speaking Our Minds*, which describes and reshapes the theoretical landscape in the evolutionary biology of communication, allowed us to compare and contrast that field of inquiry with the field of experimental pragmatics, the area we know best. Here is how the two are similar: both fields make room for the code model and a Gricean ostensive-inferential model, both recognize Grice's monumental proposal as important, and yet in each of our respective fields, it seems like a majority takes it for granted that the code model should be the point of reference.

Another generally appreciated feature of the work is the way Thom clears away misconstruals and addresses incorrect assumptions, for example when he points out that shared function does not mean shared histories: "…ostensive communication is far more expressive than coded communication, but that is no argument in favour of continuity. To argue otherwise…is also to argue…that flying evolved from walking" (§2.7), or when he refers to the qualitative difference between *code* as used in the code model among his colleagues as opposed to the way it is used as linguistic code. Interestingly, he also goes on to point out how his colleagues use the word "mean" in a descriptive way ("territory marking 'means' 'do not encroach upon this territory'") whereas Gricean formulations use "mean" in a theoretical way that addresses speakers' intentions being recognized. This allowed us to appreciate a distinction when comparing our two areas: in evolutionary biology, defenders of the code model recognize that the point of signals is to "*do* things to an audience" whereas those who implicitly adopt the code model in our area are hardly concerned with that (their focus is mostly about determining the extent to which words and grammar capture extra-logical or extra-literal meaning).

As part of our effort to draw parallels and see where concepts in his academic world line up (or not) with ours, we were however drawn to one important difference between Thom's outlook on his field and our outlook on ours – the pragmatic phenomena that are being accounted for. Whereas his explanandum is any given convention (the way a community chooses a side of the road to drive on, the

way a Pictionary game evolves, the "drift-to-the-arbitrary" in the evolution of writing symbols, finding a common language in intention-reading games, etc.), ours is on-the-spot interpretations of utterances. For instance, while we both might be concerned with metaphors, Thom is interested in the way metaphors make their way more-or-less permanently into language (§5.1), while we are interested in the way an original metaphor is processed for the first time (and *if it succeeds* in being part of communication, how does that work). His question is how do individuals manage to build conventions that render communication expressively powerful. Our question is, how are we able to say and pragmatically understand an endless number of new utterances.

The upshot is that this leaves at least a couple of places where we are no longer working in parallel. While he is interested in the process of grammaticalization, the fact that historical changes observed in languages are "overwhelmingly unidirectional" (§5.5) has not been seriously investigated from an experimental pragmatics point of view (though historical linguists are interested in pursuing this line alongside experimental pragmatists, see Grossman & Noveck, 2015). In fact, we can go further and say that experimental investigations into conventions (conceptual pacts, lexical entrainment and the like) are said to pose a challenge to a Gricean picture (e.g., see Brennan and Clark's work [Brennan & Clark, 1996], which shows how participants will use a more informative name, such as *Golden Retriever*, to be as informative as possible in a given situation but then will stick with it even later, when the more general term *dog* would do).

We have more than a passing interest in this misalignment because we would like to draw lessons from *SOM* that could have an impact in our area. However, we wonder about the extent to which that is possible. While *SOM* heralds the ostensive-inferential model, and Relevance Theory in particular, by underlining the advantages such models bring when compared to its rivals, the book – and perhaps it is inherent to the study of language evolution in general – is focused on different phenomena. While Thom's explanations are Grice-inspired, the examples seem to arrive at a nexus where Gricean explanations end and Lewisian concerns begin.

Given our shared theoretical commitments (i.e., in *SOM* and our own work), we have two reactions. One is that we suspect that there is a way to integrate Thom's framework into accounts of spontaneous pragmatic interpretation. For example, it strikes us as plausible to view coercive and cued actions as two flexible, non-overlapping categories of expressive verbal behaviour at the moment of production (before an addressee provides a reaction). Our other reaction is to remain circumspect and conclude that the answer to our query (can the approach in Thom's book be reconciled with the worries of experimental pragmatics?) is, *not readily* (and that would take nothing away from the book's brilliance). However, we remain hopeful for the former, i.e., that Thom's framework can provide the means for one to align evolutionary biological accounts of language with the everyday comprehension of utterances.

Comments

***

Thom Scott-Phillips: Towards interdisciplinary alignment?

These are forward-looking, thought-provoking comments. Thanks Ira and Tiffany for several astute observations about the similarities and differences between how evolutionary biologists and experimental pragmaticists conceive of communication. And also for a positive suggestion about how different frameworks for communication might be aligned across disciplines.

I'd like to endorse and elaborate on a couple of points that Ira and Tiffany make. First, they write that "in evolutionary biology, defenders of the code model recognize that the point of signals is to '*do* things to an audience' whereas those who implicitly adopt the code model in our area [experimental pragmatics] are hardly concerned with that". This is true, and I had actually not noticed before now quite how stark this dichotomy is! Interestingly, language evolution is an area where these two worldviews meet. Researchers working on the cultural evolution of languages, focused as they are with the origins of language structure, are not much concerned with what languages are actually used for (there are exceptions, of course). In contrast, researchers with a more biological perspective often emphasise that signals are tools for doing things to the world. What unites these two schools of thought is a commitment to the code model as a framework for linguistic communication. I hope that *SOM* will help to bring attention to the limitations of this commitment.

Second, I'd like to endorse Ira and Tiffany's observation about the possible utility of experiments on grammaticalization. I suggested in *SOM* (p. 120) that a pragmatic perspective might help to explain the directionality of grammaticalization. Experiments that test this idea will, I'm sure, face significant methodological challenges, but aside from this the fruits are rather low-hanging. Gareth Roberts has for some time being saying that historical linguists should do more experiments. It would be interesting to see what common ground Ira and Tiffany might share with Gareth.

Looking forwards, Ira and Tiffany suggest that there may be "a way to integrate Thom's framework into accounts of spontaneous pragmatic interpretation". This is clearly related to Dan's comment about the possibility of developing of general framework for communication, one that can naturally describe both the code model and the ostensive-inferential model. Ira and Tiffany are "hopeful" about the prospects for such a project. It is certainly a tantalising possibility. I don't have anything of substance to add right now, but Dan's comments and Ira and Tiffany's suggestion are giving me pause for thought.

9

ENJOYABLE, BUT DOESN'T SOLVE THE MYSTERY

By Bart De Boer

I have read Thom Scott-Phillips' book with great pleasure, but also with a very critical eye. It is extremely well written – I have read most of it during long train rides and had no difficulties concentrating on it. For someone who is as easily distracted as myself that says quite something. I also felt myself almost persuaded by the arguments, but in the end there are many things with which I disagree.

I do agree that pragmatics and inference are extremely important in language, and that they perhaps deserve more attention than they usually get. However, I also do think that Thom presents too much of a straw man argument. Most serious students of language evolution have been convinced of the importance of pragmatics and inference for a long time (although, given the complexity of the subject they may not have made it the focus of their research). The position that Thom appears to be arguing against is perhaps the syntax-centred, Chomsky-inspired formal view of language. However, this position is of decreasing importance in both linguistics and the study of language evolution. I definitely do not agree that the ostensive-inferential model proposed by Thom solves all problems of language evolution. It may help explain how communication started, but it really doesn't explain why language is the way it is (Thom's discussion of combinatorial structure notwithstanding, the ostensive-inferential model does not explain how language became combinatorial), and I was quite baffled to find nothing about acquisition in the book. I also felt extremely uncomfortable with the part where the evidence for cases where there appears to be no theory of mind but there does appear to be language was discussed. I think Thom could have been much more self-critical there and now runs the risk of falling into the pit of confirmation bias. I also think there is much more continuity between the code model of communication and the ostensive inferential model. For instance one can imagine a continuum between a pure code model, with fixed signals and meanings, and more ostensive-inferential models where meanings are progressively less well-defined and need to be acquired from context. Again, the lack of discussion of the role of acquisition is a weakness of the book. I also find that Thom overstates the importance of ostension and inference. A lot of language (and even non-linguistic communication) is extremely stereotyped. Come to think of it, Thom's favourite example of the raised and

tilted coffee cup is pretty conventionalised as well, and actually a good example of the outcome of (culturally evolved) ontogenetic ritualisation. There may be a role for ostension in establishing conventions through cultural evolution, but I am not convinced it plays a very prominent role in everyday language use. In the end, if we have learned anything in the last twenty years of studying language evolution, it is that there is no single explanation for humans having language, and language being the way it is. On the contrary, it seems clear that a number of factors have conspired, aided with a dose of coincidence, to give humans language and other (intelligent) animals not. Inference and ostension have no doubt their role to play, and I think the book makes a good point of arguing it. However, I also think the book is old school language evolution in the one-man-one-idea tradition of Pinker, Bickerton, Dunbar or Mithen, to name just a few. Such theories provide food for thought and may help to attract students to the field (in itself extremely important) but the real progress is made by painstaking and self-critical research of more detailed questions (a point also made by Richard Moore). Of course, Thom has made contributions of this nature as well, and for me his book would have been better if it had reflected this approach to language evolution more.

Comments

***

Thom Scott-Phillips: The importance of big picture thinking – and of pragmatics

I of course agree with Bart that progress in any scientific endeavour requires detailed and, yes, painstaking work. Who could doubt this? I cannot, on the other hand, agree with his disregard for big picture synthesis. He writes that "such theories provide food for thought and may help to attract students…but the *real progress* is made by painstaking and self-critical research of more detailed questions" (italics added). The implication is that big picture thinking plays little role in real scientific progress. Surely he does not really mean this?

Let me give one example of the sort of contribution that big picture thinking can make. *SOM* argues that, at the broadest level, there are two key questions for language evolution: (1) how and why did humans evolve to communicate ostensively; and (2) how and why do collections of communicative conventions emerge and then transition towards the forms that they do? I do not know if Bart would agree with this conclusion, but if it is right, then this alone represents genuine scientific progress, made possible by big picture thinking.

One other thing. Bart says he is convinced that many language evolution researchers recognise that pragmatics is important. I am myself convinced that many language evolution researchers do not recognise just how important pragmatics really is. Yes, pragmatics is recognised as relevant and important, but this recognition is mostly lip-service. One point I wanted to express in *SOM* (and this too has its roots in big picture thinking) is that if real scientific progress is to be made, then lip-service is not enough. Pragmatics should be the beating heart of the field. I'd be keen to hear whether Bart agrees with this.

10

COMMUNICATION, CULTURE, AND BIOLOGY IN THE EVOLUTION OF LANGUAGE

By Kenny Smith

Speaking Our Minds is an enjoyable book, providing an excellent survey of some of the perennial and current issues in the field of language evolution, as well as providing a clear summary of Thom's position on the central role of ostensive-inferential communication in language origins. I hope neither author will mind if I say that it reminded me very strongly of Jim Hurford's recent (and rather more monumental) books on the same topic, which is perhaps not so surprising since Thom studied with Jim. It is also worth pointing out that it is a very bold book, since some of the crucial evidence that Thom would need to nail down his position is not available – I can't help but be struck by how many cells in Table 4.2 (which in reviewing the evidence for the human-uniqueness of capacities for ostensive-inferential communication is really the heart of the book's argument) are filled "Not (yet) directly studied". I hope the book stimulates some of that work.

I want to make two comments here, on sections of the book where Thom touches most closely on work that I am most familiar with, on the cultural evolution of language (in chapter five) and (much more briefly) how that changes our understanding of what the biological capacity for language actually is (in chapter six).

The role of communication in the cultural evolution of language

Chapter five of the book outlines Thom's account of how, once the capacity for ostensive-inferential communication was in place, symbols, compositional structure and grammatical function words might emerge from processes of language learning and language use. This chapter is framed in terms of cultural attraction, which I admit was not to my taste – I think it is a pity not to acknowledge the central role that Rob Boyd and Pete Richerson have played in establishing a scientific approach to studying cultural evolution, or the value of Tim Griffiths's important contribution in putting computational models of language evolution on a much sounder theoretical footing by making them Bayesian.

But it is nonetheless a very useful, clear, coherent and up-to-date summary of the ongoing experimental work on the ways in which fundamental structural features of language arise from language learning and language use.

I was particularly intrigued by Thom's framing of the results of our 2008 paper (Kirby, Cornish, & Smith, 2008), and the emphasis on the role of communication in that work, which coincides very nicely with my own current thinking. This is all very nicely described by Thom in the book, but to recap: in that paper we describe two experiments in which participants attempt to learn an artificial language which provides labels (typed words) for objects (coloured moving shapes), and are subsequently tested on their ability to recall labels when prompted with a shape. Participants in both experiments were arranged in transmission chains, where the first participant in each chain was trained on a random set of labels (an impossible learning task), and subsequent participants in a chain were trained on the language produced during recall by the previous participant in the chain (so the language produced during recall by participant 1 becomes the target language for participant 2, the language produced during recall by participant 2 becomes the target language for participant 3, and so on). Languages change as a result of this iterated learning process. In Experiment 1, the vanilla version of the experiment, we found that the languages rapidly became *degenerate* as they were passed from person to person, rapidly shedding labels and collapsing distinctions; after 10 'generations' of transmission, the languages generally had very few labels (in the most extreme case, only 2 distinct labels to describe 27 distinct pictures). These degenerate languages are easy to learn, because they are simple, but they aren't particularly language-like because they are too simple – they wouldn't allow a user of such a language to convey many (or in extreme cases, virtually any) distinctions between objects. Of course, considerations of communicative utility aren't relevant in this experiment, because the languages are never used – they simply have to be learned and recalled, and therefore evolve to maximise their learnability, at the expense of their communicative potential. We can nicely capture this result in models which assume that learners have a prior preference for simple languages, i.e., languages with a shorter coding length; in line with established results in Bayesian iterated learning (e.g., Griffiths & Kalish, 2007), languages in transmission chains come to reflect this prior preference in learners (Kirby, Tamariz, Cornish, & Smith, 2015).

In a second experiment in our 2008 paper, we attempted to block the emergence of these degenerate languages by *filtering* the language: if a participant produced the same label for two or more objects during recall, we would pass on only one object paired with the wannabe-ambiguous label. This reduces the viability of degenerate languages, by eliminating the best evidence for degeneracy (multiple objects paired with the same label) from the data learners see. As a result, the languages in this second experiment (usually) evolved to be simple but not degenerate: they became structured, such that labels developed a compositional morphology, where parts of each label identified the colour, shape and motion of the object the label was associated with (e.g., labels for blue shapes might begin in 'l-', labels for black shapes might begin in 'ne-'; labels for bouncing shapes might end in '-plo', labels for looping shapes might end in '-pilu', and so on). In coding length terms, such compositional

languages are more complex than degenerate languages (in order to describe them, you have to write down all the morphemes in the language and the rules of their combination – in the simplest case for our experiment, this would require a list of 9 morphemes and 1 rule of combination), but substantially simpler than random languages (which can only be described by exhaustively enumerating all objects and their associated labels, in the case of our experiment requiring a clumsy dictionary of 27 entries). While these compositional languages therefore aren't the simplest possible languages (which are degenerate: to describe a degenerate language, you just have to write down a single label that can be used to label all objects), they are robust to the filtering pressure we imposed, since they provide a distinct label for every object; coincidentally, they would also be ideal for communication, allowing the maximal number of distinctions to be conveyed, although as in Experiment 1 of that paper, considerations of communicative utility aren't actually relevant in the experiment, since the participants never use the language to communicate.

We intended this filtering procedure to be a proxy for communication: rather than languages being transmitted via a process of learning and aimless recital, they are transmitted by learners learning from examples of language *use*, and we might reasonably expect language use to disfavour ambiguous signals. As Thom highlights, this experiment shows the crucial role that communication plays in the cultural evolution of linguistic structure: we only see the emergence of structured languages when there are communicative considerations at play. We have recently followed up on this work by replacing the filtering proxy with a far more satisfactory model of communication – Thom kindly cited a couple of proceedings papers outlining early stages of this work, the full version of which has now been published as Kirby et al. (2015). In that work we contrasted our Experiment 1 results from the 2008 paper (learning only, degenerate languages the result) with two conditions involving *pairs* of participants who learn a language and then use it communicate, taking it in turns to describe pictures for each other and to identify the pictures their partner describes. In the Chain condition of this newer experiment, each pair of participants attempted to learn the language produced during communication by a previous pair of participants, mimicking the real way in which languages are transmitted by learning from language use during communication, and replacing the transmission filter with actual communication. In the Closed Group condition, the same pair of participants learnt an initial random language and then communicated with each other using that language over and over: this language is therefore under pressure to be useful for communication, but under reduced pressure from language learning (the language is only learnt by naive learners once, right at the start of the experiment).

The kinds of languages that emerge in these two conditions are strikingly different, and also differ markedly from the degenerate languages, which emerge in Experiment 1 of our 2008 paper. In Chains, as one would hope, we see the emergence of structured languages; this fits the results of our filtering experiment from 2008, and shows that languages which have to be learned and used to communicate will evolve to be structured, which fits with the story Thom gives in the book. However, the languages in the Closed Group condition are different: they remain essentially holistic, being composed of a set of largely arbitrary associations between meanings and idiosyncratic signals. These holis-

tic languages are great for communication, since every object has a distinct label, but not particularly easy to learn (since they are quite complex) – but that doesn't matter in Closed Groups, because there is not much learning by naive individuals going on. Again, these experimental results can be nicely captured by simulation models in which we assume that learners have a prior preference for simple languages (i.e., languages with shorter coding length) and that learners avoid ambiguous utterances when communicating (using a simple Bayesian model provided by Frank & Goodman, 2012).

What is rather striking about this result in reference to Thom's thesis is that it shows that communication alone is not enough to produce structure – transmission to naive individuals is also required, because it imposes a pressure for simplicity, and structure only emerges when pressures for communication and simplicity are both at play. I don't actually think that is problematic for Thom's position, since he is careful to describe ostensive-inferential communication as one pressure acting on languages during their cultural evolution, but for me it is important to remember that the biases of language learners plays an equally important role.

I am also excited about the potential for these same experimental and modelling methods to address more fine-grained questions about how communication interacts with learning to shape language, which I think has the potential to speak to, and should be informed by, the issues Thom raises in this chapter and the book more generally.

Our treatment of communication in Kirby et al. (2015) is rather minimal, and probably not representative of ostensive-inferential communication in the real world: participants have to uniquely identify an object for their partner, who must pick the correct object from an array; however, the speaker doesn't know what objects the hearer's array contains, and so the best they can do is provide a label which uniquely identifies the target object among all possible objects. In other words, it offers absolutely no consideration of how context might influence the structure of languages, and in particular how ambiguity in language might be tolerated or preferred as long as it isn't detrimental to communication in context. My student James Winters is doing excellent work on his PhD looking at how reliable features of the communicative context might work their way into the structure of the linguistic system: his first paper is out now (Winters, Kirby, & Smith, 2014), although too late to make it into the book (maybe the second edition!). In the experiment reported in that paper, James trained participants on an artificial language which can be used to describe distinctive 'objects' (actually, little aliens) belonging to two categories: there are 4 distinct star-shaped aliens and 4 distinct blobby aliens. Participants then take turns describing aliens for their partner, and the language they produce during interaction becomes the target language for a fresh pair of participants, as before. However, unlike in our 2015 paper, James systematically manipulates context: the speaker's task is to produce a label which will enable the hearer to identify the correct alien from a set of two aliens, and the speaker knows the context in which the hearer will be making this selection (i.e., the speaker knows which two aliens the hearer will have to discriminate between). In one condition of the experiment, the two aliens that confront the hearer on every trial are always from the same category (i.e., two star aliens on one trial, two blob

aliens on another); in a second condition, they are always from a different category (one star and one blob in every context); in a third condition, they are sometimes from the same category and sometimes from different categories (a mix of the first two conditions).

Using this paradigm, James was able to show that the reliable features of the context in which communication takes place ends up shaping the structure of the languages that evolve through interaction and transmission. In particular, in conditions where participants are only ever confronted with contexts consisting of aliens drawn from different categories (one star, one blob), the languages tend to become partially degenerate: all 4 star aliens are associated with a single label, all 4 blob aliens with another distinct label. This language looks ambiguous when taken out of context (all the star aliens have the same label), but given the context in which communication takes place it is actually unambiguous – participants are never called upon to discriminate between star aliens or between blob aliens, so the 'ambiguity' of the language is never a problem for communication. In contrast, in the condition where participants are only ever required to discriminate within-category (i.e., differentiating between star aliens, or between blob aliens), the languages tend not to develop this ambiguity – rather, the languages retain 8 distinct labels for all 8 aliens, which allows within-category discriminations to be made. Finally, in the mixed condition where participants must sometimes differentiate within-category and sometimes between-category, the languages tend to develop an elegant structure in which the labels simultaneously encode category membership and individual identity within that category – for instance, the labels for the star aliens might be "hupa", "hepa", "hopa" and "hapa", where the '-pa' ending conveys that these are all stars (which makes discriminating star aliens from blob aliens easy) and the first syllable encodes the identity of the individual aliens (essential if you want to discriminate between a hupa and a hapa).

I think this work gives an exciting hint of how we can start to explore how communicative context shapes language, but reading Speaking Our Minds has made me consider how we need to expand our experiments and models to look at genuinely ostensive-inferential acts of communication – it may be that there are relatively minor tweaks that we can perform that will make them more informative, or it may be that we need a rather more radical rethink of how we approach communication in the lab, and what kinds of linguistic adaptations we should be seeking to explain.

Biology and culture in the evolution of language

It seems to me, based on chapter five, that Thom and I have very similar positions on how we should explain fundamental properties of design features of language (i.e., the fact that language is symbolic, combinatorial, compositional, employs ambiguity in just the right places, etc.): all these properties of language arise as a result of cultural evolution, as a consequence of people learning a language from the observable communicative behaviour of others. In Thom's thesis all these features of language therefore follow "for free" from the crucial evolutionary breakthrough, unique to humans, of the capacity for ostensive-inferential communication. Similarly, I have previously suggested (e.g., in Smith

& Kirby, 2008) that the uniqueness of human language may be due to uniquely human abilities to infer the communicative intentions of others, which is prerequisite for the cultural transmission of meaning–signal mappings; once cultural transmission of sets of such meaning–signal pairs is possible, structure inevitably follows. These two positions seem entirely compatible, and as discussed above, I like Thom's emphasis on the crucial role communication plays in shaping culturally-transmitted communication systems, which was absent from that 2008 paper.

Given this high level of agreement, I confess I was rather puzzled by sections of chapter six, in which Thom takes issue with a nice quote from Simon Kirby which I would fully endorse, that "[c]ultural transmission…provides an alternative to traditional…adaptationist explanations for the properties of human language" (Kirby, Dowman, & Griffiths, 2007, p. 5241, quoted in *SOM*, p. 136). Thoms responds that "cultural attraction does not provide an alternative to adaptationist explanations of design in nature, because this supposed contrast, between cultural evolution and natural selection, is in fact a false dichotomy" (p. 136), then goes on to describe a coevolutionary model (the model from Smith & Kirby, 2008) in which language evolution consists of two interacting evolutionary processes, cultural evolution of languages and biological evolution of the learning biases underpinning language learning. It's great to see this model discussed here, but it seems rather tangential to the central claim that, as I understand it, Simon was making in that quote: important properties of human language (e.g., the fact that language is symbolic, combinatorial, compositional, employs ambiguity in just the right places, etc.) are not biological phenomena, which need to be explained in terms of the selective advantages of these linguistic features that drive the genes ultimately coding for those linguistic features to fixation; rather, they are cultural phenomena, a consequence (through a very complex, indirect route) of the biological apparatus underpinning cultural transmission. This seems to me to be completely compatible with Thom's position, that the uniquely human adaptation behind language is the capacity for ostensive-inferential communication, and that the evolution of this capacity set in place the cultural evolutionary dynamic which lead to the emergence of languages which are symbolic, combinatorial, compositional, employ ambiguity in just the right places, etc. In other words, they both seem to agree (as do I) that we don't need to provide an evolutionary account of the fitness advantages associated with specific linguistic features, but an explanation for the evolution of the capacity for cultural transmission of communication systems in humans. In Thom's case, this boils down to providing an explanation for the evolution of the capacity for ostensive-inferential communication in humans. I might place the emphasis more on the evolution of the cognitive capacities underpinning social learning (if these are different from the capacities underpinning ostensive-inferential communication, which is something I'd like to think about more), Simon might place the emphasis elsewhere, but in all three cases there is agreement that the uniquely human biological capacity for language consists of some potentially quite high-level, abstract capacities, rather than a very detailed specification of the details of the structure of human language. We all therefore seem to be in agreement about the types of capacities, which evolutionary accounts of language have to explain; given that I'm not sure this is still a mainstream position, I think it makes sense to emphasise the commonalities between these closely-related accounts, rather than dwell on the differences of emphasis.

Comments

***

Thom Scott-Phillips: It's possible to be an adaptationist, and to think that the mapping between biology and culture is indirect

With these comments Kenny effectively updates §5.6 of *SOM*, where I discuss the role of communication in the cultural evolution of languages. He describes well the work that the Edinburgh group has been doing since I sent *SOM* to the publishers and which, he rightly surmises, I would have otherwise discussed. I have nothing of substance to add to Kenny's description here, except to confirm in particular that I would have drawn attention to James Winters' work. The way that communication is operationalised in James' experiments is an important step forward.

Since I have little to add on these empirical matters, let me turn instead to an epistemic issue that appears to divide Kenny and myself. Kenny confesses that he is "rather puzzled" as to why I took issue with Simon Kirby's assertion (with which Kenny agrees) that the findings at hand provide "an alternative to…adaptationist explanations for the properties of human language". Let me therefore try to expand on why I think it is wholly cogent to disagree with Simon (and Kenny) on this point, while agreeing with him about the empirical findings that Kenny summarises.

The key word is "adaptationist". Kenny rephrases Simon's view as "important properties of human language (e.g., the fact that language is symbolic, combinatorial, compositional, employs ambiguity in just the right places, etc.) are not biological phenomena…rather, they are cultural phenomena, a consequence…of the biological apparatus underpinning cultural transmission". This seems to strongly imply that, for Kenny (please do correct me if I have read you wrong), adaptationism includes, more-or-less by definition, a commitment to viewing properties of language as biological traits, or at least as a direct expression of biological traits (as opposed to an indirect expression, mediated by cultural transmission). My understanding is that Simon has a similar view.

Yet this is just not what adaptationism is. Adaptationism is simply a tool for studying biological traits. It entails no views one way or the other about the relationship between biology and culture. In short: adaptationists are simply committed to the view that, for a given biological trait and hypothesised function, the fact that natural selection is the only source of design in the natural world allows us to do some reverse engineering, in order to understand how the trait works, and what it is for (see §6.1). Adaptationists are in no way committed to the view that, say, writing systems, or fashions, are biological adaptations (some might believe this, but it is not a logical entailment). Same for properties of languages.

Matters are not helped here by the fact that the most prominent piece of advocacy for using adaptationism to study language – Steven Pinker and Paul Bloom's famous paper (Pinker & Bloom, 1990) – also treats languages as direct expressions of biological traits, meaning that the two views can become conflated in readers' minds. But this is not a logical entailment. Adaptationism makes no commitment as to how the relationship between biology and culture is mediated. It is possible to be an adaptationist about the biology, and to think that the relationship between that biology and cultural traits is an indirect one. Indeed, that is exactly my view, and I'd like to believe that Kenny would agree.

11

INTENDING TO SPEAK OUR MIND, AND SPEAKING OUR MIND

By Mathieu Charbonneau

Thom Scott-Phillips' contribution consists in further grounding Dan Sperber and Deirdre Wilson's Relevance Theory into an evolutionary and cognitive framework for the advent of human language. I take it that the central thesis of Scott-Phillips' book is that language is not an organ. Rather, it is the result of the joint capacity for ostensive-inferential communication – a side product of increased social capabilities for recursive mindreading – in addition to the production and cultural transmission/transformation of conventional codes. Scott-Phillips' project is a difficult one and, to me, the synthetic review he offers in his recent book is an important step forward.

I want to focus on a mechanism that, in my view, is important for Scott-Phillips' project, but not directly discussed in the book: the capacity to produce public displays from private, mental representations. I will see how it relates to some of the conceptual points made in the book – the issue about continuity and discontinuity between natural and conventional codes, and the role of natural codes in the evolution of ostensive communication.

Making intentions publicly available

Relevance theory is all about communicating (displaying and interpreting) intentions. When ostensive-communicative acts are produced, they need to signal both (1) that they are communicative, and (2) what their informative content is. However, for ostensive communication to work at all, it is not enough that one individual have clear intentions to both communicate and inform, and that the receiver be capable of recognizing these intentions when made manifest. The one communicating must also be capable of producing signals that carry with them the right kind of structure (i.e., information) so that the receiver will contextually understand what the informative content consists of, and that there is information to understand in the first place. Communicative intentions aside, the capacity to produce public representations that signal some informative content depends on (1) knowing that some signals will be relevant to another person to understand the intended content but also (2) knowing

how to produce an arbitrary signal of this kind. Having sufficient knowledge about others' minds and what sorts of signals will manipulate them in the right way is not enough for communication to be ostensive. One needs to know how to produce, from one's private mental representations, the public signals capable of transmitting the right kind of informative content (Sperber, 2006). In other words, it is of little value that one knows that some sorts of signals will allow a receiver to identify one's own communicative and informative intentions in the right way if one does not also know how to produce such public displays from the relevant mental representations and intentions. I am pointing here at a transduction problem: how do I know how to produce a public display at all (how am I capable of producing it) from my own private mental representations, and moreover, how do I make sure that the display will appropriately convey the right kind of information so that my interlocutor will be able to infer my communicative and informative intentions?

An evolutionary/neurocognitive framework requires explaining how, between ancestral and contemporary Homo forms, a specific capacity for producing language emerged. We are required to take seriously the problem of producing, with the right structure, public displays from private mental representations, especially in an arbitrary, voluntary, communicative way. Indeed, without such productive capacity, you could do all the recursive mindreading you want, and have all the communicative and informational intentions you want, you would still be unable to produce relevant communicative acts. In other words, you might be an expert at recursive mindreading, but be very poor at speaking your mind. In this sense, it seems that the productive capacity must emerge either before or at the same time as ostensive-inferential communication comes about.

Being able to solve this transduction problem is no small cognitive feat, and I am not prone to believing that the cognitive capacities to produce relevant public displays from private mental representations come for free with the capacity for recursive mindreading.

Consider the case of expressive aphasia, also known as Broca's aphasia. Expressive aphasia is a form of agrammatism: the inability to produce grammatically correct sentences. There are diverse degrees of severity to this handicap, with the most severe case consisting in only being able to produce a single word for all communication acts, as Paul Broca's well known patient Tan was only capable of. Expressive aphasics usually are capable of comprehending others, so in the context of ostensive-inferential communication, they seem to be able to recognize the communicative and informative intentions of others. What they lack is part of the capacity to produce public signals that convey such intentions: they fail to properly map the sounds they emit on the intended meaning they wish to convey. This does not mean they cannot communicate at all in other ways, but it does show that producing meaningful acts of communication is not a trivial task.

The dysfunctional mechanisms in the expressive aphasia case are likely to have evolved in order to produce more powerful ostensive communication, and thus after ostensive inferential communication was already in place. But this only pushes back the problem as to which productive processes were in place to allow the evolution of ostensive-inferential communication in the first place. Indeed, even

an early ostensive-inferential communication system faces the same problem, i.e., individuals must still possess the capacity to produce – in the right way – public signals that reflect their private mental intentions.

Pointing as a natural code

So what would this process look like in an early ostensive-inferential communication system? I suggest here that some natural code, namely pointing, might serve as a key candidate for such early ostensive-inferential productive mechanisms. Moreover, I would suggest that pointing has opened the way for ostensive-inferential communication to evolve through a process of ritualization (§2.5, Table 2.2).

Consider the following, admittedly crude just-so story for the evolution of pointing behaviours by ritualization. Infants incapable of grasping objects out of their reach will tend to stretch their arm in the direction of the objects they desire, in a proto-pointing manner. Such behaviour, which serves first as an attempt to grasp the object, "signals" part of the mental states of the infant: the infant has some intentions about the object it is grasping towards. Later on, parents could come to recognize the intention of the infant by associating the direction of the kids reaching behaviour with the desired object, and recognize that the infant desires the object being "pointed" at, and so learn to offer the object to the infant. We do not have ostensive-inferential communication here yet, because the infant is just trying to grasp the object, not making his intentions public to others in any intentional manner. Nevertheless, such behaviours can fixate and serve as a natural code for wanting/asking for some specific object, as long as parents react appropriately.

In this speculative scenario, pointing serves as a natural code as it is based on a pair of associations (one between the desired object and the grasping/pointing behaviour, another between the grasping/pointing behaviour and the parent's response). The scenario looks a lot like how chimpanzees learn to point. Indeed, Leavens & Hopkins (1998) and Leavens, Hopkins, & Bard (2005) argue that captive chimpanzees learn to point at desirable food that is out of their reach in order to manipulate a nearby human in bringing the food to them. However, as Scott-Phillips argues (with Tomasello [2008]), chimpanzees do not have sufficient mindreading capabilities to communicate ostensively. Pointing in chimpanzees must thus be a natural code. In contrast, for humans it seems that such processes are ontogenetically well entrenched as little nurturing is required for human infants to actually produce such reaching/pointing behaviours (Kita, 2003; Tomasello, 2008). Moreover, humans (even infants) use pointing as a conventional code. They do not only request objects, they also use pointing to inform others cooperatively (Kita, 2003; Tomasello, 2008).

If the capacity of pointing in our ancestors was first used as a natural code as the scenario above suggests, we would have a story to explain how recursive mindreading might have co-opted pointing as an early productive mechanism. However, that would imply that pointing started as a natural code

and became a conventional code once ostensive communication had evolved. This does not mean that language conventions, specifically, evolved from pointing as a natural code. Rather, it means that pointing made ostensive communication possible by serving as an early productive mechanism, and then became a conventional code. How can this be possible?

First, note that pointing behaviours (in chimpanzees for instance) are much more expressive than the set of natural codes that Scott-Phillips considers in his criticism (§1.5, §2.7). I agree that there is just so much we can express with sneezing, frowning, etc., even when we use these behaviours in an ostensive way. However, pointing, in contrast to sneezing, frowning, etc., is a much more expressive natural code as it is a natural way to publicly display part of what our mind is privately thinking of (its intentions). Part of the content of your mental states is directly made public as its referent is directly being pointed at. In an early, pre-ostensive context, pointing directly shows what the object of a private mental intention is, and thus can serve as an early form of producing public displays of mental representations in an expressive way. I would suggest that it is this expressive power of pointing, even as a natural code, which makes pointing a plausible productive mechanism that could have served as a platform for the evolutionary acquisition of ostensive communication.

Moreover, from this natural code, and once ostensive-inferential capacities have evolved into the picture, not only the infant can communicate her intention to reach an object (imperative), but with further development of mindreading and cooperation, the pointing act can then play expressive and informative roles (Tomasello, 2008, p. 111–143). In fact, with pointing plus ostensive communication, we have a very expressive form of communication that does not obey the code model anymore. For instance, pointing in one direction can mean many things: we can refer to the chair being pointed at, its color, its shape, its number, or to the person that usually sits on it, etc., the specific intended meaning changing with the context (Wittgenstein, 1953/2001, §33). This certainly is not language, nor is it as expressively powerful as language, but it is much more expressively powerful than sneezing ostensively.

Scaffolding and transformative continuity

Scott-Phillips recognizes pointing as a key mechanism for early ostensive communication (§5.2). This is in line with Tomasello's work (who also discusses pantomiming as an early form of communication in his 2008 book) that pointing in addition to words can make an expressive proto-language as it allows for the production of combinatorial signals (section 6.1 in Tomasello, 2008). This would also be in line with Scott-Phillips' view that ostensive-inferential communication allows for a more direct route to communication. However, as I have speculated above, pointing might have first evolved as a natural code and only then been co-opted as a conventional code. This puts some stress on Scott-Phillips' arguments against the continuity of natural codes with conventional codes, or at least it suggests that there are different kinds of continuity between natural and conventional codes that might fit into Scott-Phillips' scenario about the evolution of language.

Scott-Phillips makes a good point that it is ostension and inference augmented with conventional codes that makes language possible, not natural codes augmented with ostension and inference (§2.7). Pointing behaviours have not evolved into words, I agree. Nevertheless, pointing plausibly has served as a preliminary step in the evolution of ostensive communication by solving the basic productive problem discussed above. Moreover, pointing might have served as the early platform for the evolution of combinatorial communication with conventional codes (see Tomasello, 2008, for arguments supporting this last point). What these two roles point at is the usefulness of pointing as a simple productive mechanism for communication. So there seems to be a form of continuity here between natural codes and language, assuming we read continuity as how we got from A to B, i.e., how a natural code (pointing) helped scaffolding the evolution of ostensive communication (of which language is a special case) by serving as an early productive mechanism.

In contrast to the scaffolding understanding of continuity suggested here, Scott-Phillips understands continuity as how A transformed into B, where B is a modified version of A. More specifically, Scott-Phillips argues that natural codes did not (and could not) evolve into languages along some gradient of expressivity (p. 48; see also Origgi & Sperber, 2000). He convincingly argues that a transformative understanding of continuity is wrong, showing that conventional codes are not natural codes plus pragmatics, and that this argues against the code model. I agree. However, where I would disagree with Scott-Phillips is when he appears to infer from this rejection of the transformative view of continuity that natural codes should be removed from the evolutionary picture of the advent of language. I admit, this conclusion is not stated explicitly. However, it is suggested by the way natural codes are dealt with throughout the book. For instance, there are no more references to natural codes in Scott-Phillips' book once chimpanzees are shown to be unable of ostensive-inferential communication and only capable of using natural codes. This suggests that natural codes were last useful in our closest common ancestors, but not afterwards, thus being a key difference-maker in the evolution of the human lineage's capacity for language.

Continuity understood as a process of evolutionary scaffolding does not entail the transformative view. When I say that pointing might have evolved from a natural code to a conventional code, I do not mean that language as such is just pointing plus pragmatics (see above). What I am suggesting is that pointing was originally a natural code, serving first as a pre-ostensive mechanism of expressing private mental representations in the form of public signals. When ostensive communication came along, pointing was co-opted as a conventional code, retaining its role as a productive mechanism along the way, but then served the additional role of a conventional code.

Perhaps I am over-reading Scott-Phillips here. This "implicit dismissal" of natural codes as relevant to the evolution of language – e.g., as relevant/important evolutionary steps in the origins of language in modern humans – seems to me to be founded on a false dichotomy, i.e., that either you are a transformative continuist (what Scott-Phillips blames the code model to be), or you are a discontinuist (Scott-Phillips). I believe that the scenario elaborated above, a form of scaffolded continuism, offers a third

alternative. Ostensive communication (of which language is a special case) might have evolved from natural codes, not as a transformation of natural codes, but by co-opting natural codes (i.e., pointing) as a productive mechanism for expressing informative intentions. The scaffolding reading of continuity offered here in fact would agree that "linguistic communication is made possible by the existence of [natural] codes", at least phylogenetically. However, it would not agree that because natural codes are part of the evolutionary story of ostensive communication – and thus language –, it means that language functions according to the code model. Rather, it means that natural codes – specifically pointing – might have served as the necessary productive mechanism required for the capacity of ostensive communication to evolve in the first place.

If the arguments presented here are anywhere close to being sound, then they put stress on some aspects of Scott-Phillips' arguments. First, I have argued that an important condition for ostensive communication is the evolutionary acquisition of a capacity for producing public displays from private mental representations. An evolutionary and cognitive account of ostensive communication (and thus language) depends on explaining how such productive mechanisms have evolved. Scott-Phillips does not explicitly discuss the evolution of such productive mechanisms. I hope I have offered reasons convincing enough to see that we should take the evolution of productive mechanisms seriously for a complete theory of the evolution of ostensive communication, and thus language. Second, I have argued that even when pointing serves as a natural code (e.g., as in chimpanzees), it directly expresses part of the mental states of the pointer, and that pointing is a much more expressive and flexible natural code than what Scott-Phillips suggests natural codes can be, even without ostensive communication. This makes pointing a plausible candidate for an early productive mechanism scaffolding – with the addition of recursive mindreading – the evolution of ostensive communication. Moreover, I have suggested that a natural code serving a productive role can be co-opted by ostensive communication and become a conventional code. This does not mean that language works according to the code model. Rather, it means that natural codes can have a role to play in the evolution of ostensive communication, and thus that there can be a form of continuity between natural codes and conventional codes.

Comments

***

Hal Morris: Reply to Mathieu

I'm mostly in agreement with Charbonneau's comments.

One caveat – pointing arising as described by Charbonneau, is an example of ontogenetic ritualization – a naive attempt to accomplish something (like a dog pushing on a door to try to open it), or an infant trying to push its way to the mother's breast, or pushing the mother's back down to allow climbing on, gets attenuated and ritualised into a signal – the dog's pawing the door as a request for you to open it, etc. Scott-Phillips and I think that Tomasello even more emphatically (2008) describes ape gestural com-

munications as being deliberate and conventional – in contrast to vocalizations, which Tomasello states are involuntary responses to some state of agitation.

Deacon in the *Symbolic Species* (1997), mounts a strong argument about the difficulty of achieving the complex vocalizations, as breathing is primarily involuntary, so to emit a long and complex string of sounds requires exceptional breath control, and so unsurprizingly occurs in some sea mammals, and birds (which I believe must synchronize breathing with flyling motions since unsynchronized inhaling and exhaling with their capacious lungs messes with the airfoil). Where Deacon, I think, was just pointing out a huge barrier to get over, I would, with Tomasello, see it as a very strong argument (along with how naturally the deaf take to sign language) for complex communication starting out in gestural mode.

The mastery of pointing, along with closer attention to direction of gaze, provides one gesture that does the work of a thousand words – there being so many things one could meaningfully point at, and if pointing and mindreading (or better yet, joint intentionality per Tomasello) coevolved, they would tremendously complement each other, making the evolutionary pay-off for the former more commensurate with the difficulty of achieving it. It is also plausible that, especially with pointing in place, signs for group tactics, as in hunting would evolve more easily than signs for other uses, such as naming various things. It might even make possible a scene such as this description of a hunting party in New Guinea by E. Richard Sorenson, with only that level of repertoire:

> "One day, deep within the forest, Agaso, then about 13 years of age, found himself with a rare good shot at a cuscus in a nearby tree. But he only had inferior arrows. Without the slightest comment or solicitation, the straightest, sharpest arrow of the group moved so swiftly and so stealthily straight into his hand, I could not see from whence it came. At that same moment, Karako, seeing that the shot would be improved by pulling on a twig to gently move an obstructing branch, was without a word already doing so, in perfect synchrony with Agaso's drawing of the bow, i.e., just fast enough to fully clear Agaso's aim by millimeters at the moment his bow was fully drawn, just slow enough not to spook the cuscus. Agaso, knowing this would be the case made no effort to lean to side for an unobstructed shot, or to even slightly shift his stance. Usumu similarly synchronized into the action stream, without even watching Agaso draw his bow, began moving up the tree a fraction of a second before the bowstring twanged."

You can find the source of this citation here.

Such an example seems to me to speak strongly for group intentionality (the citer of the passage spoke of "group proprioception"). Otherwise, who is doing what to whom?

Thom Scott-Phillips: A welcome elaboration

I think there are genuinely new insights here. In fact, I think they are sufficiently substantial that I'd like to encourage Mathieu, if he are so inclined, to work them up into a short journal article.

Readers might be surprised (or maybe not?) that I agree with a great deal of what Mathieu writes. In particular, I agree (i) that how mental representations are turned into ostensive signals is understudied and not well understood, and (ii) that pointing could very plausibly provide a way to scaffold the use of conventional codes. Mathieu suggests that I might disagree with him at least on the second point, but he also astutely observes that *SOM* does not actually do this explicitly. In fact, his proposals seem very plausible to me. In what's below, I clarify and nuance this agreement.

As for (i): clearly mental representations are turned into public expressions all the time (for instance, whenever I use a tool, or, more basically, whenever I act in a goal-directed way). Moreover, such behaviour is widely studied in, say, cognitive psychology and cognitive neuroscience. What is less studied and less well-understood are those cases where the translation of mental into pu-

blic is done in the name of influencing others. As Mathieu puts it, "how do I make sure that the display will appropriately convey the right kind of information so that my interlocutor will be able to infer my communicative and informative intentions?" This is the production side of communication, and pragmatics, and indeed the philosophy of language on which it draws, is much more focused on the comprehension side of things: how do listeners overcome the problem of underdeterminacy? Relevance Theory states a computational solution to this problem, and uses that to derive an algorithmic explanation too (I did not enter into these details in *SOM*). It does not, however, do the same for production, and I did not offer any solutions in *SOM* either. Mathieu is right to point out that there is much work to be done here.

As for (ii), Mathieu outlines how "recursive mindreading might have co-opted pointing as an early productive mechanism". I do not see it as quite the challenge to my views that Mathieu himself does. He infers that I likely believe that natural codes "should be removed from the evolutionary picture of the advent of language", but immediately acknowledges that "this conclusion is not stated explicitly". In fact, I had not considered Mathieu's suggestions when I wrote *SOM*, and in light of his arguments I am persuaded of their plausibility. To me, his suggestions fit naturally into §5.2 and §5.3, where I discuss the beginnings of ostensive communication systems. I explicitly suggest that icons and indices may be important ways to bootstrap the first symbols. I could and perhaps even should have added that existing natural codes may be an important source of such iconicity. Mathieu further adds the potentially important point that there is at least one natural code that is not only iconic, but is also expressive of a mental state: pointing. He hence believes that pointing is "a plausible candidate for an early productive mechanism scaffolding – with the addition of recursive mindreading – the evolution of ostensive communication". He may indeed be right. However, I see this not as a challenge to the thesis of *SOM*, but as a welcome elaboration of it. I would like to thank Mathieu for this, and encourage him to develop this point further.

12

ONE EXPLANATION TO RULE THEM ALL?

By Clark Barrett

The field of language evolution, it seems to me, is a microcosm of the evolutionary behavioural sciences more generally, in the following sense: you can maintain more or less any position you want, even in the face of data. Is there a Universal Grammar? Some are convinced there is and others are equally positive there isn't, with subjective probabilities for the two hypotheses hovering in the high nineties and low single digits, respectively, in the opposing camps. Is spoken language evolutionarily old, or relatively recent? Take your pick. Are there language-specific cognitive adaptations, or not? It depends on your postal code.

Amidst this free-for-all, many have tried their hand at finding the holy grail of language evolution: the single unique feature from which all the rest of language's notorious complexity follows, the one explanatory ring to rule them all. Examples include Michael Tomasello's candidate, shared intentionality, and Hauser, Chomsky, and Fitch's recursion. With *Speaking Our Minds*, Thom Scott-Phillips introduces and defends his own candidate for what makes human language special: ostensive-inferential communication, which is in turn made possible by recursive mindreading. *SOM* mounts an impressive theoretical argument, and along the way makes a strong plea for the importance of bringing pragmatics to the fore in thinking about language evolution. I think the praise that the book has received is well-deserved, and I can't imagine any serious scholar in the field of language evolution won't feel compelled to read it and to engage with its arguments.

My comments about the book can be arranged into two basic categories, in descending order of positivity. First and most importantly, I could not be more enthusiastic about *SOM*'s emphasis on the importance of theory of mind and pragmatics in understanding language evolution. I agree that theory of mind is likely to have played a much larger role than anyone has yet recognized in enabling not only linguistic communication, but cultural transmission and cooperation more generally. To me the importance of mindreading in enabling and stabilizing many forms of human sociality has been criminally underexplored, and I'm not sure why. One possible explanation is Tooby and Cosmides' notion of "instinct blindness": mindreading underlies so much of everyday social interaction, and we do it so

effortlessly, that we scarcely notice its operation or feel it necessary to invoke it in explaining communication and cognition. Whatever the reason, we are still largely in the dark about how important various forms of mindreading might be in solving the social adaptive problems that must be solved to make human communication and cooperation stable. Thom and I are entirely on the same page about this.

My second category of comment takes a more skeptical turn. While I find Thom's account of the uniqueness of human language plausible, I am by no means convinced that he has identified, in recursive mindreading and ostensive communication, the prime causal mover that gave rise to the rest of human linguistic complexity. Because I take a broad view of what mindreading is – and I believe that basic components like intention-reading are phylogenetically widespread –, I assume that mindreading, broadly speaking, far predates language. Moreover, I agree that some fairly sophisticated mindreading abilities such as belief tracking had to be in place before important aspects of human language, such as implicature, evolved. However, I am not yet convinced that multi-level recursive mindreading is the main element that gave rise to the rest of linguistic complexity. It's certainly possible, but I don't think the evidence is yet sufficient to call the game for recursive mindreading. Indeed, I think there are several links in the causal story offered in *SOM* that, while plausible, still remain to be verified. Thus, while I think the book's focus on the importance of mindreading and pragmatics in language evolution is spot-on, I think it takes a victory lap too soon in declaring the major mysteries of language evolution solved.

This is true for both theoretical and empirical reasons. Empirically, I think we just don't have enough evidence to say for sure whether many of the book's key claims are true. And theoretically, I think there remain some mysteries even in some of the more basic mechanisms that Thom calls "well understood". In the epilogue at the end of the book titled "The Big Questions Answered", Thom's ninth and final question and answer are:

> Q: "Coherence: Does the proposed account depend only on well-understood evolutionary mechanisms, or is it more speculative?"
>
> A: "My proposals depend on well-understood evolutionary mechanisms alone."

I suppose one man's understanding can be another man's confusion, and I might be the confused one here. But as *SOM* points out, linguistic communication depends on solutions to deep problems of cooperation – in particular, the problem of what makes linguistic communication honest – and I don't think the underlying mechanisms are well understood at all. Thom opts for reputation as the stabilizing mechanism, and it's true that there exist game theoretic models in which reputational costs stabilize signal honesty. In that sense, the mechanisms in those models may be "well understood", but I don't think that's equivalent to saying that the mechanisms that *actually* stabilize cooperative communication in human language are well understood. Indeed, problems of stability in linguistic communication are a subset of problems of large-scale cooperation more generally. While it's clear that these

problems *have* been at least partly solved in human cooperation – for example, I can engage in successful communication with a stranger I'll never meet again – it's not at all clear how they have been. Indeed, the problem I attributed to the field of language evolution, i.e., that you can believe more or less anything you like, seems to apply just as much if not more so to the field of the evolution of cooperation. Claims of "X" and "not X" coexist quite stably in this literature, as seen in debates over group selection, strong reciprocity, and the like. So, I don't think we'll be ready anytime soon to check the "Question Answered" box for what stabilizes large-scale cooperation.

Then there is the claim that what makes human linguistic communication special is ostensive-inferential communication, which in turn depends on recursive mindreading. Here again, while I think there is a plausible causal story, I'd like to see both better theoretical support in the form of evolutionary models, and much better empirical support from work in humans and other animals, including evidence from everyday linguistic communication outside the lab. It's clear, from work by Thom and others, that humans *can* do recursive mindreading. *SOM* also offers some tentative evidence that only humans can do this, and that only humans have ostensive-inferential communication. However, showing that humans have these abilities and that chimps do not is not the same as showing that these abilities causally enabled the evolution of human linguistic complexity. Indeed, that is difficult to show, because recursive mindreading and ostensive communication are offered as the ultimate causes of human linguistic complexity, and causation is notoriously difficult and perhaps impossible to show using the comparative method. This is especially true when there is only one taxon that shows all of the traits in question (us). And these aren't the only cognitive traits that are uniquely derived in us; there have been changes in brain size, executive control, planning, tool use, social learning, social complexity, and more. Which factor is causal, if any? Moreover, while recursive mindreading of many levels can be shown in the lab, we don't yet know how much everyday speech depends on many-level recursive embeddings of mental state inference, nor how much recursive mindreading is actually predicted by the theory. What amount, or lack thereof, would falsify the theory? Or is the mere demonstration that humans can do it enough to call the question settled? Because this is a theory that depends on a complex causal cascade, many of the steps of which have not yet been fully theoretically elaborated or empirically demonstrated, I think we have a long way to go before declaring all of the Big Questions answered.

To me, though, that's no reason not to take seriously *SOM*'s argument for the likely importance of mindreading in language evolution, and to do so without considering the Big Questions of language evolution settled. Indeed, in my view, we're better off proceeding without declaring a winner, because I doubt that there is one true explanatory ring to find. Instead, my hunch is that mindreading has been just one factor among many in enabling the gene-culture coevolutionary pathway leading to the current state of human linguistic complexity. If that's so, the answers to the Big Questions will come not in the form of simple propositions involving easily stated concepts such as "ostensive communication", but rather, a messy causal graph including many directed edges that we have yet to discover or name. Still, that's no reason not to turn our attention right now to the importance of mindreading in lan-

guage evolution and cultural evolution more generally. I hope that *Speaking Our Minds* convinces others, including researchers and granting agencies, that this is a quest well worth pursuing.

Comments

***

Thom Scott-Phillips: Yes, there is much more research to be done

"The field of language evolution, it seems to me, is a microcosm of the evolutionary behavioural sciences more generally, in the following sense: you can maintain more or less any position you want, even in the face of data." This is a pessimistic note on which to begin. It's the sort of complaint one expects to hear from critics of evolutionary approaches. Clark, however, is in fact a leading advocate and practitioner of evolutionary thinking in the human behavioural sciences (his recent book, *The Shape Of Thought* [2015], is an unapologetic synthesis of some of the most prominent lines of evolutionary thinking in the human behavioural sciences). I am therefore a little surprised by his pessimism, and I can't agree that, in the evolutionary behavioural sciences in general, one can get away with any view at all, even in the face of data. Look at, for instance, research on human attraction, attractiveness, informed by sexual selection theory, where for the past 20 or so years, hypotheses and conclusions have been continually updated in the face of data.

Having said that, I do think Clark is right to sound at least a slight note of pessimism about language evolution in particular. Hypotheses, if not quite free, are pretty cheap especially when it comes to the biological evolution of the relevant cognitive abilities. One significant reason for the scientific free-for-all is that before we even get to evolutionary issues, there is much debate and indeed acrimony about what the relevant cognitive abilities even are. More generally, I think that language evolution is insufficiently integrated with the rest of the evolutionary behavioural sciences. Clark does not say this himself, but I'm pretty sure he'd agree.

Anyway, on the substance. Clark thinks that *SOM* draws its conclusions too fast, and declares "victory" too soon. I certainly did not mean for *SOM* to give the impression that all is solved and there is nothing more to see here, but I acknowledge that Clark is not the first person in this book club to interpret it in something like those terms. Perhaps a future edition of the book will qualify and nuance things a little more (having your book critiqued by so many informed people is a sure fire way to make you want to rewrite significant parts of it!). But, as I say, declaring everything resolved was not the goal. Instead, what I wanted to do was to describe a basic framework – for which I do think there is a great deal of support – that can contextualise the many and various unresolved issues, and also to offer some solutions to the most important and pressing of those issues. Although apparently sympathetic to *SOM*'s proposals, Clark wants more arguments and more data to support them. Fair enough. Obviously I would welcome any and all additional data of relevance. I get the impression that, for the most part, Clark and I differ just on how much confidence we should place in the conclusions drawn in *SOM*.

Having said all that, I think Clark is in places a little unfair on *SOM*. The clearest example is his comments on the evolutionary stability of human communication. After summarising my answers, Clark writes that "problems of stability in linguistic communication are a subset of problems of large-scale cooperation more generally", and then points out that we won't "be ready anytime soon to check the 'Question Answered' box for what stabilizes large-scale cooperation". This is true, but *SOM* never attempted to explain cooperation as a whole. It only suggested that the subset of such problems, those associated with ostensive communication, are straightforwardly solved by repetitional effects (and Clark doesn't disagree with this conclusion). It's not reasonable to criticise *SOM* for not also answering bigger questions than the ones it aims to. That's outside the remit. Still, these disagreements over weight of evidence are quibbles in the bigger scheme of things. I infer from these comments, and from personal correspondence, that Clark believes that the ideas expressed in *SOM* are broadly on the right track. If his goal with these comments is to make a plea for more relevant theory and data, he will find no disagreement here.

13

INFERENTIAL COMMUNICATION AND INFORMATION THEORY

By Greg Bryant

Speaking Our Minds is a timely book that very effectively frames many of the current important problems facing researchers interested in the nature of language and communication. Too few scholars today are worried simultaneously about evolutionary psychology and pragmatics, and ever since my introduction to Thom's work with his article "Defining biological communication" I have found myself often in high agreement with his conclusions. That said, the main point of contention I would like to bring to the group here is actually a fairly fundamental issue in which my own view has recently shifted away from modern pragmatics theory. The issue concerns the purported distinction between the code model of communication and ostensive-inferential communication, and what it really adds to our current understanding. Is the supposed failure of a code model based on information theory an outdated (and false) argument from a previous time in pragmatics that needed a straw man? I'm thinking it is.

One concern at the heart of the code model problem involves the role of inference in communication. Unconscious inference is a central concept in cognitive science: it's everywhere. Even the most low-level adaptive problems in perception involving feature detection and integration are solved through inferential procedures. Bottom-up information across modalities is often structured statistically but highly impoverished, so computational solutions incorporate rich priors to extract meaningful data that can be further processed, eventually leading to perceptual experience. Top-down processes are central to all of cognition, and importantly, communication. I'm sure Thom would agree. There does not seem to be much objection to the application of information theory to these topics, so it's not just a problem about inference on unobservable data. But then what is it exactly? If we are to approach human communication generally, and linguistic communication specifically, from a computational point of view, it is hard to imagine how one might model these inferential processes free from information theory.

The primary problem as Thom describes it, following Sperber and Wilson (1986/1995), is that the code model and ostensive-inferential communication require different internal mechanisms to

function. A coding scheme involves associations where encoded symbols are sent through a channel with noise and then decoded, ideally recovered with perfect resolution but often only probabilistically retrieved. The content of the message is physically in the signal, as opposed to ostensive-inferential communication where senders provide evidence of a meaning that must be inferred. But here is where an important misinterpretation might exist: the precise way the signal is associated with the content of the message (the coding scheme) is unspecified in information theory. Shannon's original formulation, of course, was not designed to explain language understanding, or even human communication for that matter, but instead was provided as a highly general framework describing the formal problem of information transmission in noisy systems. The specific computational nature of a given communication problem depends on the given task demands. As Thom well describes throughout his book, linguistic conventions, including grammar and lexical semantics, likely evolved in the context of an existing ostensive-inferential cognitive environment, so it seems completely reasonable to assume that the coding constituted by linguistic input has been shaped by the inferential abilities of decoders. The generality of the information theory framework can accommodate this circumstance. According to the now standard view in pragmatics, once inferential meaning must be generated based on decoded evidence, the code model reportedly fails. A closer look at the formal properties of information theory does not make it obvious why this should be so. If we view the derivation of some coded message as a potential state change in a receiver, there is no rule about the particular input-to-output relationship, such as an isomorphism in surface features, or exactly how the content manifests itself in the code at all – only that the receiver's state has a possible reduction in uncertainty as a function of decoding some aspect of the signal. In a given communication system, the structured input by a sender can be tailored specifically for the receiver by design, and the acquisition of that message with the associated reduction of entropy (relative to a state without that signal) can manifest itself in any number of ways and with any number of alternatives. As Thom points out, code models can involve inference. The problem is not in a limitation of information theory – the problem is we don't know exactly how inferential communication actually works.

The beauty of inferential models of communication comes from the amazing insight that "arbitrary" symbol systems are used strategically by people in conjunction with a number of signalling channels to help targeted listeners derive relevant meanings in context-specific ways. At some level with language use, the code will likely involve a duality of patterning that affords a retrieval of the specific linguistic content, but the decoding story doesn't have to end there. It's not clear formally why a distinction between natural codes and conventionalized codes prevent a computational solution that involves the probabilistic recognition of implicit meaning via the relevant processing of meaningful symbols as well as other continuous data streams. Principles of relevance describe how senders will signal, through multiple simultaneous channels, not only communicative intentions, but informative intentions. A code model can handle that if there are coevolved encoding and decoding algorithms: encryption is a common application of information theory. In the present case, the system is designed to maximize relevance, but proximately that can look a lot like encryption, and sometimes actually is (e.g., Clark & Schaefer, 1987).

Donaldson-Matasci, Bergstrom, & Lachmann (2013) used information theory to quantify relationships between noisy environmental cues and effective developmental strategies, and a similar analysis could be applied to how listeners infer implicit meaning through linguistic evidence. One important difference being that the sender is providing a designed signal, rather than an organism processing an environmental cue. Imagine a sender providing linguistic input, which has coded structural features, with the processing outcome measurable as entropy. This just involves the words – the sender's meaning is not directly observable. In relevance theory terms, the exact words provide evidence of speaker meaning. In information theoretic terms, the literal words constitute a linked process that allows for a measure of conditional entropy. The question then becomes: how much is entropy reduced in an unobservable process (speaker meaning) as a function of observing a linked process (sentence meaning)? Another way of putting it would be to say that the cognitive effect of some ostensive communicative act is the receiver's uncertainty regarding the implied message prior to receiving the signal minus her posterior uncertainty after receiving it. One advantage here is that the formal model affords a quantification of cognitive effects, something relevance theory technically lacks. Of course, a full computational account would need to specify the various sources of information that play into comprehension algorithms (e.g., nonverbal signals, contextual details, common ground, etc.), but in theory, these are potentially explicable and quantifiable processes that critically involve a code in information theoretic terms. The bottom line is this: if interlocutors are able to reliably use given structure in proximal signals to reduce uncertainty about unobservable meanings, they must be reliant on some set of mutual representations. The arbitrariness or novelty doesn't really matter if the code is relevant and has predictable cognitive effects.

Finally, it is worth noting here that at some level, if one is to accept the proposition that brain activity is instantiated as a complex system of adaptive neural coding schemes, then it is hard to get around a code model for communication that is rooted in information theory. Communication behaviours are implemented in neural systems in the brain and body, and as Thom rightly points out, involve adaptations for both the production and perception of signals. Relevance theory, in combination with an evolutionarily informed information-based approach, and a developed theory of cultural transmission, provides the tools for a comprehensive theory of communication and cognition.

I want to thank Thom for some earlier discussion on this (I don't believe the point I'm raising is a surprise), Ray Gibbs, and in particular to Clark Barrett for bringing this issue to my attention longer ago than I care to admit.

Comments

***

Olivier Morin: Will the Code Model please stand up?

Let me thank Greg Bryant for this post, and allow me to try and dissociate two issues. One is the scope of information theory, the other the usefulness of the code model. Once we dissociate the two, we might be able to see why information theory can be useful without reintroducing a code model of communication.

I agree that interpretation, like all other cognitive activities, is an activity that reduces some uncertainty about the world. In other words, it falls under the scope of information theory. What I fail to see is how that vindicates the code model in any way. In fact, I see the association between information theory and coded communication as almost coincidental. Shannon happened to develop his theory while working on information transmission, but (as Greg Bryant himself notes) that theory is both much more than a theory of communication, and much less than that. Not to put too fine a point on it, this quote from Warren Weaver (in his introduction to Shannon's treatise) spelled things out quite clearly:

> "The word information, in this theory, is used in a special sense that must not be confused with its ordinary usage. In particular, information must not be confused with meaning. In fact, two messages, one of which is heavily loaded with meaning and the other of which is pure nonsense, can be exactly equivalent, from the present viewpoint, as regards information. It is this, undoubtedly, that Shannon means when he says that the semantic aspects of communication are irrelevant to the engineering aspects...The concept of information developed in this theory at first seems disappointing and bizarre – disappointing because it has nothing to do with meaning, and bizarre because it deals not with a single message but rather with the statistical character of a whole ensemble of messages…"

Now, if information theory is nothing more or less than a general theory of the representation of uncertain events, providing a way of quantifying the computational constraints that weigh on any such representation, there is no reason to wed it to a code model of communication. Shannon happened to be working on coded messages, but that was just a historical accident (he could have been, say, building a computer to record casino draws, or doing theoretical work on thermodynamics).

I am quite persuaded that one could, as Greg Bryant points out, have an information-theoretic take on relevance and interpretation (if only we could quantify the information carried by gestures, words, mutually manifest context, etc.). Does this vindicate the code model? It depends on what one means by it. The trouble is, once we put information theory aside, it is not very easy, for me at least, to grasp what the code model actually consists in, if it is anything more than a foil. In fact, as I think about it more, the differences between code-model and ostensive communication seem to boil down to just one thing: the presence or absence of mindreading.

I find it hard to deny that code-model communication may involve complex inferences: as Thom points out (p. 7), decoding can be a highly non-trivial task. As Thom also insists, code-model communication may also be used intentionally and strategically. So rule these out as difference-makers. And (Thom again) code-model communication can be non-deterministic, and may involve a lot of uncertainty. Rule that one out too. I am sure Thom would agree that research on animal codes shows coded communication to be context-sensitive as well. Bees don't follow any danced signal to any place whatsoever: they use their knowledge of the terrain to react appropriately.

So, if I could be so bold, I would say both the ostensive view and a properly understood code model should agree that neither complex inferences, intentional and strategic use, flexibility, or context-sensitivity set ostensive communication apart.

I may be misunderstanding something here, but I find it hard to square Thom's definition of code-model communication as depending on "mechanisms of association" alone (p. 5–7) with his insistence that code-model communication can be inferential, flexible, probabilistic, used strategically, etc. With a loose enough definition of "mechanisms of association", we can probably fit anything in, but then I don't see how we can escape the conclusion that ostensive communication is exactly similar, in that regard, to code-model communication. After all, you could always say that it simply "associates" signals with speakers' meanings, albeit in a flexible, probabilistic, context-sensitive (etc.) way.

The one thing that, it seems to me, sets ostensive communication apart is, then, the use it makes of mindreading capacities. Not everyone here might share this view, though.

Thom Scott-Phillips: On code and inferenc

Greg's comments are challenge to a fundamental part of the thesis of *SOM*, and indeed to Relevance Theory itself (and some other pragmatic frameworks).

Following Relevance Theory, *SOM* makes much of the distinction between code model communication and ostensive-inferential communication. Greg asks whether this difference is as real as it is presented. After all, all cognitive inference is information-theoretic at some level. Olivier is right to point out, in response, that information theory and the code model are not the same, even if information theory is often presented as the canonical illustration of the code model (as it is in *SOM*). Still, Olivier shares some of Greg's worries: "I don't see how we can escape the conclusion that ostensive communication is exactly similar…to code-model communication".

Here is what *Relevance: Communication and Cognition* has to say: "Inferential and decoding processes are quite different. An inferential process starts from a set of premises and results in a set of conclusions which follow logically from, or are at least warranted by, the premises. A decoding process starts from a signal and results in the recovery of a message which is associated to the signal by an underlying code. In general, conclusions are not associated to their premises by a code, and signals do not warrant the messages they convey." (Sperber & Wilson, 1986, p. 12–13)

Here is another way to think about it: If ostensive communication can in general be described in terms of a code, then what is the code? What is the "underlying code" that links sticking my tongue out with the sentiment that these people are all idiots (see *SOM*, p. 7)? There is of course no such thing, and hence this is not decoding. Comprehension is instead a matter of inference: your conclusion that these people are all idiots is warranted by the premises that I make manifest when I stick my tongue out. Greg might respond: "Ok, fair enough – but on this notion of inference, aren't many supposed cases of code model communication actually inferential?" Quite possibly many cases of communication involve inference in some way. Non-human primate communication almost certainly does. But the point – the dividing line between coded communication and ostensive communication – is not whether inferential processes are involved, but whether inferential processes are sufficient to make communication possible in the first place.

Let me expand. One form of inference of particular importance for communication is metapsychology. It is important because, when developed to a sufficiently rich degree, it allows communication to take place even in the absence of any code. This is ostensive-inferential communication, and this is what I mean when I say that human communication is made possible by mechanisms of metapsychology (I do not think that I expressed this point quite so clearly in *SOM*). Other systems are made possible by codes, and can be made more expressively by inferential processes. In some cases, in particular non-human primate communication, those processes might even amount to metapsychology. That does not, however, amount to ostensive communication.

Dan Sperber: Inference is ubiquitous, ostension is very special

Thanks to Greg for his post and to Olivier and Thom for their comments. This further comment is intended to be relevant to the issues they all three discuss.

When Deirdre and I introduced the notion of ostensive-inferential communication, the "ostensive" bit referred to what the communicator does and the inferential referred to what the audience does. This, we have progressively realised, may be misleading. Inference is ubiquitous: in perception, in memory, in motor control, in fact in all aspects of cognition, including coding–decoding and the production of ostensive stimuli. We had in mind, as Thom rightly stresses, just the very special kind of metapsychological inference done by the audience from the communicator's "productions" (i.e., perceptible behaviour or perceptible traces of behaviour) to the communicator's communicative intention (in which the informative intention is embedded). We now prefer to talk (not yet reflected in our publications) just of ostensive communication, leaving the inferential part to the gloss.

This should help make it clear (with the caveat that any possible misconstrual is going to occur) that we are not denying (and are obviously not committed to denying) that the use of a code can, in principle, be cognitively quite complex, involve rich inference, and even, indeed, as Olivier points out, strategic (that is, involving metarepresentations of what other agents may think or intend including about what we may think or intend). None of this, I believe blurs the distinction with ostensive communication, which is not defined by metarepresentational complexity, not even by their strategic character, but by what these metarepresentations are about (namely, communicative and informative intentions). Also, let's remember that, unlike human languages, which would be grossly defective as tools for coding–decoding communication, and are adapted to serve a major but nevertheless subordinate role in ostensive communication where they render the range of what can be communicated limitless, the evolved codes of animal communication are adapted to the communication of very narrow ranges of information of here-and-now relevance. As far as I know, there isn't even compelling evidence of an intended strategic use of signals in animal communication.

On the other hand, as Csibra and Gergely have pointed out, ostension permits the transmission of general knowledge. I surmise that only ostension does.

Greg Bryant: What's in a code?

I think one important issue has to do with how we define "code". I am using the word in a very general sense – a sense I believe is relevant to information theory. A code is any system of rules converting information into physical signals (numbers, letters, sounds, gestures, etc.) which then affords a computational process by which that information can be used to reduce uncertainty in a receiver. It is true, as Olivier points out, that the code model is separate from information theory, but I would argue that they seem to be used interchangeably by Thom, as well as by Sperber and Wilson and others. I maintain that we might want to put the code model straw man to bed, and instead focus on how information theory can help us address empirically the computational problem of ostensive communication.

Thom writes: "An inferential process starts from a set of premises and results in a set of conclusions which follow logically from, or are at least warranted by, the premises. A decoding process starts from a signal and results in the recovery of a message which is associated to the signal by an underlying code. In general, conclusions are not associated to their premises by a code, and signals do not warrant the messages they convey."

Unless you limit the definition of coding quite severely, I would argue that deriving a conclusion from a set of premises is easily construed as a form of decoding. The question is in the details of how the premises are represented by structured signals and what sort of algorithmic processes constitute the logic of the derivation from those signals. This is essentially what I meant earlier when I mentioned that the limitation in our understanding is rooted in our lack of specific knowledge about how inferential communication actually works. So in describing a decoding process as recovering a message that is "associated" with the signal, I'm not

seeing any specific conflict with the description of inference. Both are sufficiently vague to be mutually compatible when understood in an information theory framework. As I understand Olivier, he makes a similar point.

Thom then writes, "What is the 'underlying code' that links sticking my tongue out with the sentiment that these people are all idiots (see *SOM*, p. 7)? There is of course no such thing, and hence this is not decoding." This reveals what I believe is a limitation in Thom's definition of coding. I would argue that there is some relationship between the structural features of the expressive action and the likely inferences a target receiver will draw, and this relationship can be reasonably construed as a coding system of some sort. The details of that relationship aside (which, of course, are not currently understood, hence our problem), the physical signal systematically reduces uncertainty in perceivers regarding the contents of the message.

Finally, Thom states: "But the point – the dividing line between coded communication and ostensive communication – is not whether inferential processes are involved, but whether inferential processes are sufficient to make communication possible in the first place." I agree, but doesn't this distinction break down if the coding scheme is evolved to exploit inferential capabilities?

Dan makes the important point that ostensive communication is about a particular kind of metarepresentation. I think there are good reasons to believe that the mechanisms generating these kinds of representations have a unique and identifiable developmental trajectory, and have some species-specific features, but there are also phylogenetically related systems we can observe in other animals as discussed by Katja in her commentary. I was curious, however, about a statement made by Dan at the end of his comment: "As far as I know, there isn't even compelling evidence of an intended strategic use of signals in animal communication." I am wondering what exactly qualifies in his view because to my mind there are innumerable examples of this in the nonhuman literature. Perhaps it depends on one's use of the word "strategic"? Functional deception in many animals (e.g., primate and bird species) satisfies the bill for me, including even chickens, not an animal with the most revered cognitive abilities. To take the chicken example (e.g., Marler, Dufty, & Pickert, 1986), few, if any, would argue that when producing false food calls in the presence of strange females, roosters are trying to strategically change the mental state of the hens. Rather selection has likely shaped calling behaviour by altering triggering mechanisms that increase copulation frequency, and thus reproductive fitness. But in other cases, such as the use of food calls in strategic ways by capuchin monkeys (e.g., Di Bitetti, 2005), I believe there are very solid reasons for assuming the mental states of other monkeys are part of the computational system, regardless of whether the monkeys are conscious of it or not.

To categorize a communicative act as ostensive must there be an explicit awareness of the specific strategic mental state manipulation? A similar question is also being considered in current debates on figurative language understanding (e.g., Gibbs, 2012). I would say no.

Dan Sperber: *Intended* strategic use of signals in animals

Greg writes: "I was curious about a statement made by Dan at the end of his comment: 'As far as I know, there isn't even compelling evidence of an intended strategic use of signals in animal communication.' I am wondering what exactly qualifies in his view because to my mind there are innumerable examples of this in the nonhuman literature."

I am well-aware of the examples of deceptive behaviour, including, less frequently, deceptive use of signals in the animal literature. I should have not just said but stressed that I had in mind "intended strategic uses", where the agent intentionally engages in one course of action rather than another as a function of other agents' possible responses. To not just do this but intend to do this, you need to be able to represent – consciously or unconsciously, this isn't the issue – how what you do influences what other might do in response. There is no compelling evidence that I know of that shows that non-human animals do this in their signalling behaviour.

Di Bitetti, whose interesting work on delayed food call among capuchin monkeys Greg cites, does not provide clear evidence of strategic thinking. Di Bitetti himself cautiously concludes: "finders seem to withhold the production of food-associated calls under certain conditions in a functionally deceptive way". Meaning, I take it, that intentional deception isn't established.

Greg Bryant: One more example

Dan is being a better skeptic than I am. I also agree that the literature on deception in non-humans, unfortunately, often ignores signalling issues. I don't intend to play the game where Dan smacks down various empirical examples, but one more example comes to mind that directly addresses Dan's requirement that "the agent intentionally engages in one course of action rather than another as a function of other agents' possible responses."

Santos, Nissen, & Ferrugia (2006) show that free-ranging rhesus monkeys, when faced with two containers of food – one that can be opened silently and one that can only be opened by ringing an attached bell – will preferentially choose the silent one when the researcher is averting her gaze, but will choose randomly when being viewed. These researchers definitely interpret the monkeys' behaviour as being strategically designed to manipulate others' knowledge states, and they have similar work examining how eye gaze provides a cue to knowledge states. Of course, corvids are also well documented changing their caching behaviour as a function of what conspecifics can and cannot see.

I am sympathetic to the skeptical stance here, and when I teach this stuff, this is the challenge I pass on to my students. But my instincts tell me that the problem here is mostly methodological.

Dan Sperber: Deception – yes. Intentionally deceptive signalling – no.

Greg gives good examples of deception in monkeys and corvids. They might even involve an intention to deceive. The deception however isn't done by means of signalling. What I would find much more surprising (and, hence, extremely interesting) would be clear cases where animals intentionally use signals in order to misinform.

Greg Bryant: Vocal control is a problem

The trick, of course, is showing it "clearly". As Dan points out, demonstrating the capacity to alter behaviour as a function of different possible responses from target audiences shows the underlying psychology needed – I feel like this is the real reasoning hurdle. Rhesus monkeys have been shown to have some volitional control over their vocalizations (Hage, Gavrilov, & Nieder, 2013), so maybe it is just a matter of time before they are documented vocalizing in ways that manipulate knowledge states of others. Perhaps another major limitation has to do with vocal and gestural control, in addition to mindreading abilities.

14

NATURAL LANGUAGE AND THE LANGUAGE OF THOUGHT

By Deirdre Wilson

I found Thom's book extremely illuminating, insightful and enjoyable. I learned a great deal from it, and look forward to this discussion, from which I am sure I will learn a lot more.

One point where I was left feeling rather frustrated was in the brief discussion of Chomsky's views on language and adaptation (§6.2). I had been hoping to get some guidance on how to think about the increasingly acrimonious debates between Chomsky and others on the existence or non-existence of a dedicated language faculty or Universal Grammar, but Thom remains officially neutral on this. As he says in the précis,

> "What might be a natural object of study is an innate cognitive mechanism – sometimes called a Universal Grammar – without which we would not be able to acquire and use languages. I say that this only 'might be' a natural object of study simply because whether such a mechanism actually exists is a disputed and much vexed issue, on which I am personally agnostic."

Although I would have liked to hear more on the pros as well as the cons of Universal Grammar, what mainly frustrates me is the possibility that what Chomsky means by language is not the same as what Thom means, so the discussion may be at least partly at cross-purposes.

Thom makes a convincing case that language evolved to make ostensive communication expressively powerful, whereas Chomsky repeatedly denies that language has a primarily communicative function. Thom defines language as "[t]he suite of cognitive traits that allow us to acquire and use languages" (i.e., public languages like French or English). For Chomsky, though, language seems to be more like a language of thought, and this has become increasingly obvious in his recent writings. Here are a couple of extracts from a recent informal talk by Chomsky on 'Language and the Cognitive Sciences' at Carleton University (my italics):

> "It appears overwhelmingly clear that a generative process suddenly emerged at some pretty recent point... Well it emerged in an individual, mutations don't take place in groups, so some

individual was fortunate or unfortunate enough to get this generative capacity… Furthermore, there was no selectional pressure at that time. There couldn't be. It's just something that happened to an individual. So you'd presumably expect what appeared at that point to be just determined by natural law, there's no other pressure, something kind of like a snowflake. And the same would be true as this capacity of this property is transmitted to offspring. Notice that the capacity itself HAS selective advantage, the person who had this capacity *could think, it could plan, it could interpret, you know, could construct, internally of course, complex thoughts*. That, you would expect, would have advantage transmitted to offspring in some small hunter-gatherer group, maybe a couple of hundred people. It could take over most of the group after some period, and at that point there would be a reason to externalise it, to make it available to others so that it's not just in your own head. And that seems to be the way language works, *with externalisation being an ancillary process*."

And here is a longer extract from the questions after that talk:

"*Question*: One of the things you argued in support of is Language's productivity, and Fodor uses the same argument for a language of thought. I wondered if you think language plays a greater role, like Fodor does, and is there such a thing as a Universal Grammar of thought?

Noam Chomsky: Well, I DO think so, and, in fact, what I was describing was a language of thought. What actually seems to have happened, as far as we can piece anything together, and as far as the empirical evidence shows, is that at some point – maybe seventy-five thousand years ago – some small neural rewiring took place, of course, in some individual, because it's the only possibility, and that individual had a *computational process which was somehow linked to pre-existing conceptual structures*. Now, what those pre-existing conceptual structures are, we haven't a clue about that. That's the problem I ended with. Nobody has a clue what that could be or where it could have come from. But it's there, and it's totally different from anything in the animal world. *And if this generative procedure could link to it, you do have thought. So, that's a language of thought. And then somewhere down the line it got externalised* and you get interactions among individuals. *So, yeah, I think it is a language of thought. But I don't see any reason to think it's separate from language, I think it just is language*."

Chomsky goes on to add, rather cryptically:

"*Noam Chomsky*: In fact, if you look at Fodor's work, and you ask the question 'What do we know about the language of thought?', well it turns out to be English.

[Audience murmurs]

Question: You are saying that everybody thinks in English.

Noam Chomsky: I'm not saying that we think in English, but the reason it turns out to be English is that's the language everybody's using. *Whatever the language of thought – this internal, our own internal language, yours and mine – whatever it is, is inaccessible to introspection*. Okay, if you introspect, you can't go one minute without talking to yourself. It takes a tremendous act of will not to talk to yourself. In fact you do it all night, it keeps you up all night. It's just impossible not to do it, but what you're introspecting is the externalised language. So, you can tell when you're talking to yourself, you can tell whether two sentences rhyme, okay. Or you can tell how long they are, or something like that. And actually, if you really pay attention, you're not really talking to yourself in sentences, just kind of odd little fragments. *Something is going on deeper which we CAN'T introspect into, any more than you introspect into the mechanism of vision, and that's the language of thought. And it's probably universal. It's hard to imagine how it could be anything else. There's no evidence for acquiring it."*

Now if this is what Chomsky means by 'language', it's easy to see why he denies that language has a primarily communicative function and that it emerged through natural selection. On the other hand, as long as the capacity for ostensive communication was present before "externalisation" took place, Chomsky's account might well be compatible with the view that the *externalisation* process was driven by selective pressures, along just the lines Thom suggests. (I don't know if this makes sense – it's a long time since I thought of myself as a linguist – and I am quite happy to be put right.)

Comments

***

Thom Scott-Phillips: On Chomsky's evolutionary proposals

Deirdre asks about Chomsky's views on language and adaptation. Before I answer, let me take this opportunity to thank Deirdre for her work in developing Relevance Theory. As I expressed in SOM, I believe that Relevance Theory provides a genuine scientific paradigm for pragmatics, and for this work alone, all of us interested in cognition and culture owe her a huge intellectual debt.

Deirdre observes that I made very few comments about Chomsky's views on language evolution in SOM. What I did argue (see p. 134–136) was that if there is a cognitive mechanism worthy of the name Universal Grammar, then it would have to have been selected in order to aid the acquisition and use of (what I call) languages i.e., sets of conventions created in order to enhance the expressive capacity of ostensive communication.

As Deirdre observes, Chomsky would be unlikely to endorse this terminology. In line with the quotations that Deirdre picks up on, Chomsky defines language as "a computational cognitive mechanism that has hierarchical syntactic structure at its core" (see p. 1 of the paper that Deirdre links to). The presently preferred hypothesis about what this mechanism consists of is a very simple one: "human language syntax can be characterized via a single operation that takes exactly two (syntactic) elements a and b and puts them together to form the set {a, b}" (ibid.). This operation is called 'merge'.

Let us grant Chomsky all of the above. Let us accept for the sake of argument that this "generative procedure" is "a language of thought"; that, rather than being separate from language, this procedure "just is language"; and that all that communication brings

to the table is externalisation. Deirdre suggests that this view could be compatible with the one I put forward in SOM: "as long as the capacity for ostensive communication was present before 'externalisation' took place, Chomsky's account might well be compatible with the view that the externalisation process was driven by selective pressures, along just the lines Thom suggests". And she is right: the two views are compatible in just this way.

This acknowledgement of compatibility should not, however, be taken as an endorsement of the Chomskyan approach to evolution. Chomsky believes language, thus defined, is uniquely human. To the extent that I understand what 'merge' is, this seems unlikely to me. 'Merge' is not computationally difficult, and the computation of putting things together seems necessary for all sorts of animal cognition. Perhaps I have misunderstood the idea, but if I have not, it seems to me that much more elaboration is required to support the claim that 'merge' is uniquely human. (It is possible, however, that it was only in humans that there was selective pressure for such an ability to be externalised in communication. Indeed, this is exactly the point I made about Universal Grammar in SOM [see above].)

A second issue is adaptation. If 'merge' is uniquely human, then that invites the question: Why? Why should such a trait have evolved only in one species, and why this species? Adaptationism can provide candidate hypotheses to such questions, but Chomsky does not offer any. On the contrary, he is famously hostile to adaptationism, and hence refuses to even propose such an explanation. He instead suggests that the emergence of language is, effectively, just a lucky accident ("what appeared at that point to be just determined by natural law, there's no other pressure, something kind of like a snowflake"). To be blunt, this seems to me an abdication of intellectual responsibility. I agree with Steve Pinker (among others) that if one is committed, as Chomsky is, to studying cognitive computation like one would any other biological trait, then adaptationism is a critical part of the toolkit.

Nicholas Allott: More on Chomsky on the evolution of the language faculty

I hope it's alright if I join in with this very interesting discussion, even though I haven't yet read Thom's book.

A few brief points about Chomsky's views on the evolution of the language faculty. First, Tim Wharton in his comment makes a crucial point: it's not at all clear that languages are sets of conventions. Obviously there are different theories about what conventions are, but I can't see that any of them would make conventional (e.g.) the putative fact that grammatical relations are mediated by C-command. So Chomsky certainly wouldn't – and shouldn't – agree with Thom that "if there is a cognitive mechanism worthy of the name Universal Grammar, then it would have to have been selected in order to aid the acquisition and use of (what I call) languages i.e., sets of conventions created in order to enhance the expressive capacity of ostensive communication". And so, as Deirdre says, there is a real risk that Thom and Chomsky are simply talking past each other. Thom's suggestion in his reply to Tim Wharton that all the syntactic differences between languages are conventional even if shared properties of languages are not, is interesting, but not much more plausible, I think, especially given the extra commitment that they be "created in order to enhance the expressive capacity of ostensive communication". At any rate, one wants to know how the idea would go in detail.

I also agree with Deirdre that disagreement about 'adaptationism' is inessential to Thom and Chomsky's disagreement. Chomsky's proposals with colleagues about the evolution of core of the language faculty can be stated in a way that is neutral about such arguments among evolutionary biologists, I think. Chomsky, Hauser and Fitch make the bold conjecture that FLN (the 'narrow faculty of language'), which is defined as what is specific both to humans and to language, is just the operation 'merge', which puts together any two linguistic constituents, plus the requirement that the structures generated are legible by the two systems that need to use them, the conceptual-intentional (i.e., thought) system and the sensori-motor (i.e., phonetic) system. As a slogan: "interfaces + recursion = (the core of) language".

Part of the motivation is to explain what Chomsky and co-authors take to be the narrow time-frame between the appearance of natural language in the species and its dispersal over the planet, since which time it apparently hasn't evolved (given that infants from everywhere can acquire any language with apparently equal facility). A related motivation is to dispense with what earlier

syntactic theory (Government-and-Binding-era generativism) seemed to require: the simultaneous evolution of several distinct 'modules' of the grammar – theta-theory, Case theory, etc. – which would be largely useless without each other.

Anyway, as Deirdre says, the proposal is that a single evolutionary event in an individual gave him/her the ability to put together atomic thought components recursively. This ability spread to a group. I think that it's inessential to the account whether it was selected for, or just not selected out, or was a side-effect of some other change that was selected for. After this there was another event that allowed it to be externalised, and this has not been selected out – again, for whatever reason. If adaptationists like the overall story, then they are free to claim that selection pressure was crucial – following the initial mutation, of course – at both stages.

Thom doubts that 'merge' is uniquely human, saying that “'[m]erge' is not computationally difficult, and the computation of putting things together seems necessary for all sorts of animal cognition". Maybe. But do we have good evidence of unboundedly recursive cognition in other animals? That seems doubtful. Since that is what 'merge' provides (given that it can operate on its own output), then they either don't have 'merge' or some other limitation prevents them from using it in the powerful way that we do.

One final note: I see no reason to suppose that "all that communication brings to the table is externalisation" and I doubt that Chomsky thinks this. After all, he often suggests that we have pragmatic competence, which we draw on along with our linguistic competence in (some of) our use of language. Equally, selection pressures connected to communication might have been important for the development of parts of FLB (the 'broad faculty of language'), that is, the pre-existing systems and abilities that became part of a full language capacity with the emergence of FLN. Certain properties of the articulatory apparatus are pretty good candidates, for example.

15

CATS, TACS, AND KUNVENSHUNS

By Tim Wharton

I agree with Thom completely that pragmatics has been under-represented in discussions of the evolution of language (with the notable exceptions Thom mentions). I was, I recall, the only pragmaticist speaking at Evolang in Paris in 2000.

As someone whose interest in relevance theory has come via linguistics, rather than say, psychology, anthropology or cognitive science, I will not address the areas of the book with which I broadly agree – the centrality of ostensive-inferential communication, the emergence of language as a tool to make that more explicit, mindreading, cultural attractors, etc. Much of the book is, as far as I can see, right. However, there is one thing I'd like to take issue with.

On page 19 Thom introduces two pieces of terminology from my 2003 Mind and Language article (Wharton, 2003) and 2009 book (Wharton, 2009): 'natural codes' and 'conventional codes'. These, he admits, he has adapted to suit his own purposes, but the definitions he offers are in pretty much the same spirit as mine. 'Natural' codes are codes such as those used in communication that relies on strict coding and decoding (bee-dancing, bull-frog calls, etc.). In the class of natural codes I include human behaviours such as smiling, with the proviso that these can be recruited for use in ostensive-inferential communication by being either deliberately shown (in the Gricean sense) or even faked.

Conventional codes, on the other hand, are those regularities perpetuated by tacit agreement between members of a particular community: driving on the left (or, inexplicably, the right…); Morse code; leaving a gratuity in a restaurant; the person who initiated a phone-call calling back in the event that you are both cut off. Thom then goes on to define language as "the rich, structured collection of conventional codes that augment ostensive-inferential communication within a given community" (p. 20).

But here is where we differ. You see, I presented the notions of 'natural' and 'conventional' codes not in order to point out that language is an example of the latter, but rather to point out that the human linguistic code is, crucially, *neither*.

Let me explain: I could just about be persuaded that we, as members of the same speech community, could all agree to call 'cats' 'tacs', or even spell 'convention' <k-u-n-v-e-n-s-h-u-n>. Indeed, something along these lines goes on when young people decide as a group (and to the exclusion of old fuddy-duddies such as myself) to describe a positive experience as 'really bad' or a really cool band as 'really sick' (or, for that matter spell the word 'cool' <k-o-o-l>).

But there are properties of language – the headedness of phrases, the fact that dependencies are local, so-called 'island' effects and many more – that surely *cannot* be the result of tacit agreement among members of a speech community. Moreover, these properties of language cannot be induced by children acquiring that language because they are simply not there in the data they hear.

I'll quote William Lycan, as I do in my 2003 paper:

> "…most sentences of a language are never tokened at all; since hearers instantly understand novel sentences, this cannot be in virtue of pre-established conventions or expectations directed on those sentences individually." (Lycan, 1991, p. 84)

So, what I'd like to ask is this: How does 'language' as defined as a set of conventional codes fit with the notion of language viewed from an internalist, modular, domain-specific perspective, or, for that matter, with the myriad advances made by generative linguists? Relevance theory, as I understand it, was conceived as a framework intended to complement such a view of language.

And relatedly, to touch on some points raised in chapter six, to what extent does the view sketched in *Speaking Our Minds* really still allow room for some innate specification in language evolution? Thom doesn't appear to rule out the possibility of an evolved Universal Grammar-like mental faculty, but in the end he sits on the fence. Then, on page 136 he proposes that the cultural attractor account is an *alternative* to Universal Grammar. This makes me feel uneasy and I'd like to know more.

Comments

***

Thom Scott-Phillips: What is a convention?

It was Tim who first drew attention to the distinction between natural codes and conventional codes, and greater awareness of the difference between these would without question represent progress in language evolution. Perhaps *SOM* will help to facilitate this progress.

In the comments here Tim suggests, contrary to the view I put forward in *SOM*, that (at least some parts of) languages cannot be conventional codes: "there are properties of language – the headedness of phrases, the fact that dependencies are local, so-called 'island' effects and many more – that surely cannot be the result of tacit agreement among members of a speech community". Why not? Tim endorses William Lycan's elaboration of the problem: "most sentences of a language are never tokened at all; since hearers instantly understand novel sentences, this cannot be in virtue of pre-established conventions or expectations directed on those

sentences individually". Perhaps an important first point to make is that, contrary to this quote, it's not sentences that are conventions, it's the component parts, including the structural elements. This includes words, phonological patterns, morphosyntactic operations, and so on. Critically, all of these are tokened. It's in this way that languages are sets of conventions.

But I suspect this is missing the point. Tim may have something else in mind. He draws attention to some textbook example's universal, or near-universal patterns in the world's languages (headedness, local dependencies, island effects), and comments: "these properties of language cannot be induced by children acquiring that language because they are simply not there in the data they hear". This is a classically Chomskyan, poverty of the stimulus argument.

But who says these are properties of language? They are properties of languages, but we cannot assume that the properties of languages derive directly from a cognitive mechanism that we would willingly call language. Just as, say, systems of kinship cannot be assumed to be direct manifestations of a faculty of kinship, we cannot assume that language structure is a direct manifestation of a faculty of language. General cognitive constraints may suffice. Moreover, culture is not simply the design of the mind writ large. So the question is: Are these properties of languages the causal consequence of a faculty of language worthy of the name, or are they the causal consequence of less domain-specific aspects of the mind?

This latter hypothesis is certainly plausible. Arguably the most substantial empirical finding of the field of language evolution to date is one that I summarised in §5.7 of *SOM*, that semantic compositionality can be explained in large part by two factors that are not domain-specific cognitive mechanisms: expressivity (languages should be good for communication), and learnability (languages should be as easy to learn as possible). It is certainly possible that there is a similar story for all properties of languages. See, for instance, the examples I listed in a preview of *SOM* that I wrote for Replicated Typo. Regarding the sort of properties that Tim draws attention to, one hypothesis would be that learners modify languages in the direction of forms that are easier to memorise, in which case ease of memorability would be an important factor of attraction (this is just a speculative hypothesis to illustrate the point). Whatever the relevant factors, the important point to make here is that in the process of acquiring a language, individuals sign up to conventions, and sometimes contribute to changing the conventions at the same time. In fact, I'd say this is true by definition: acquiring a language just is learning a community's given set of conventions (with perhaps some intelligible modifications), and tacitly agreeing to abide by them.

Where does this leave Chomskyan Universal Grammar? In §5.7 of *SOM* I argued that the agenda for research into the cultural evolution of language should focus on the following question: what are the relevant factors of attraction for each linguistic feature of interest? As I said above, for one feature (semantic compositionality), we have already identified two critical factors, and for other features, some other factors have been identified. What I leave open is the possibility that a Chomskyan Universal Grammar is an important factor of attraction in some cases. At the same time, it may be the case that there is no such Universal Grammar, and that none of the factors of attraction that influence language evolution are language-specific. It is in this way that the cultural evolution view is a possible alternative to Chomskyan Universal Grammar.

Dan Sperber: On conventions

I find the issue of conventionality a difficult one, both for conceptual and for theoretical reasons. Here I just want to make two points in haste.

Agreeing with the insightful ideas that the cultural evolution of languages adjust them to evolved psychological dispositions leaves open the issue of whether (or better: the degree to which) these dispositions are language specific. For the sake of comparison, take right-handedness, an evolved feature of the species and a potent factor of cultural attraction. Many cultural practices, shaking hand, cleanliness usages, placing of cutlery, and so on, have culturally evolved in line with this statistical regularity. Could the advantage in coordination provided by a side bias have been a factor in the evolution of this bias? In principle yes (this is by way of illustation, I am not claiming that such is the case). In language too, it is possible – and some of Chomsky's linguistic arguments

are highly relevant here – that some of the psychological dispositions that are factors of attraction in the cultural evolution of languages evolved biologically to influence the cultural evolution of languages.

In thinking about the notion of convention, I find Ruth Millikan's 1998 paper (Millikan, 1998) more on the right track than David Lewis's and related views.

Tim Wharton: Some more thoughts

Firstly, yes, as you say, sentences are not conventions. But if language is a conventional code, the structures internal to those sentences (read utterances) – including, as you also say, phonology, morpho-syntax, etc. – are conventional, and tacitly agreed on by language users. Those structures, you point out, "are all tokened". But they're not: children are not exposed to every linguistic structure; despite this they acquire them all. That's why it's taken linguists so long to fathom out the kind of universals I mentioned. I thank Dan for pointing me in the direction of Millikan's paper (1998) on conventions: I will revisit it.

Secondly, Thom writes of the language universals I identify: "But who says these are properties of language? They are properties of languages, but we cannot assume that the properties of languages derive directly from a cognitive mechanism that we would willingly call language." I think there is some arguing past each other going on here, and return to a point Deirdre made. I say these have to be properties of language because language is the cognitive mechanism. The object of study is not languages, but language. And this, I submit, is much, much more than a mere terminological quibble. Accounts that criticize Chomsky's views on the evolution of language nearly always mistake what he means by language.

Here's a nice joke from Georges Rey, who writes some interesting stuff in this area:

> "A linguist asked him [Chomsky]: 'Noam, I like your new work, but you can say the following in Welsh' [and he produced a Welsh sentence that would have been excluded on Chomsky's view]. To which Chomsky replied: 'Well, that just shows you Welsh is not a natural language. In fact, come to think of it, it's commonly presumed that people speak natural languages. There's not a shred of reason to believe it.'"

Finally, and returning once more to your response to Deirdre, I think you oversimplify 'merge'. Yes, 'merge' is indeed the recursive operation that works on syntactic objects, but 'merge' comes in different varieties: 'set merge', 'list merge', 'pair merge'. And as well as 'merge', there are the general principles that are required to tell you to project the information in a certain way (which gives you headedness), and the cyclicity that leads to island effects and more. I don't know of any other species that has anything like the human ability to put things together the way 'merge', together with these general principles, puts things together.

David Adger: Principles of computation too?

I think that Thom, and many others, underestimate the problems in deriving universal properties of language as a cognitive mechanism from externalities interacting with general (but innate) principles of learnability and communication following Simon Kirby's work.

For example, bound variable anaphora (things like 'every child thinks he deserves a present', with the covariant reading of the pronoun, where the quantifier binds the variable), are subject to conditions across languages (whenever the phenomenal lay of the land lets it show), where the covariant reading tracks the scope of the quantifier, and the scope of the quantifier is determined by the finiteness of the clause that contains it. The poverty of stimulus issue here is particularly sharp, but even without it, the basic empirical issues are, I think, already decisive. The best (in fact only empirically successful) accounts of this all need to refer to principles that, while they may be at play elsewhere in cognition, are not principles of communication or learnability at all (I have a recent paper on *lingbuzz* here about this particular issue that's currently being revised, so any comments welcome). They are principles of structure generation and periodicity, and chunking of interpretation, etc. That is, they are basically principles of the

internal manipulation of configurations (in the case at hand, linguistic configurations), and their interpretations.

I couldn't really care less whether these were language specific (so it may be that the set of language specific non-data-derived properties of human language could be zero, so no Universal Grammar in the technical sense of the term), but they are, I think, deep principles of certain aspects of cognition. My guess is that these principles interact with equally general principles of learnability, memory, processing, social structure, etc., to give rise to the panoply of phenomena we see. But without them, there's not a snowball's chance in hell of explaining relative clauses, parasitic gaps, bound variable anaphora, scopal properties of modals and negation, constraints on subject extraction – and basically the huge literature of phenomena discovered over the years in generative grammar works. You can see this by looking at non-generative frameworks, like cognitive or construction grammar, that, when they analyse these linguistic phenomena posit structures of quite immense complexity that are just stated to be extra-linguistic, but which, actually still require a great deal of specificity, it's just specificity that's tied to, say, the terms of Gestalt psychology.

Now, there is a certain amount of nostra culpa here, as we generative grammarians haven't been good enough in explaining, in accessible terms, the nature and robustness of our results, but nevertheless, these results tell us about crucial properties of languages that reveal deep regularities in language (the underlying mechanism), and they are very robust.

Thom Scott-Phillips: Reply to David

I actually agree that learnability and communication are not the only factors at play here. (Simon Kirby might disagree – I don't know.) Indeed, I said as much in SOM: "Clearly, numerous attractors are important for the cultural evolution of languages. I identified two above, for the purposes of exposition…but there will be many more" (p. 124). (There is – mea culpa – a small mistake here; I should have written: "Clearly, numerous factors of attraction are important for the cultural evolution of languages". Still, I think my meaning was clear nevertheless.) So I totally agree that deep principles of cognition are critical. The challenge is to identify exactly what these principles are. As such, one important subsequent question is: are these principles language-specific? (David may not care what the answer to this question is, but many people do, including many generativists.)

David Adger: Reply to Thom

I guess I'd assumed that the attractors Thom were discussing weren't the kind of computational principles I had in mind, but I'm probably wrong about that. Thom says that the challenge is to identify exactly what these principles are. But I gave a few in my comment and in general I think we have a pretty good idea about at least some of these principles (principles of constituency formation, of interpretive domain, of structure alteration, etc.). These principles may not be language specific, in that they may be used elsewhere in cognition (e.g., music, arithmetic, etc.), but it's pretty hard to explain various phenomena of human language without them and our best understanding of them comes from investigation of syntax/semantics across languages. So I guess the question for Thom would be, is something like 'merge', or say cyclic interpretation of structure, a possible attractor in Thom's view?

16

A CLOSER LOOK AT COMMUNICATION AMONG OUR CLOSEST RELATIVES

By Katja Liebal

As a researcher interested in the gestural and facial communication of great apes, I want to offer some comments and facts from a comparative perspective on human communication and language evolution.

As a more general comment, I specifically liked the way Thom Scott-Phillips navigates the reader through this book by providing definitions of different terms, particularly of those often causing confusion when being used by scholars of different disciplines, as well as summaries of the most important facts of each chapter together with a brief outlook about what the reader can expect in the following chapter. Furthermore, although I am interested in potential precursors of human language in other primates, thus favouring a continuous approach to language evolution, I agree that researchers interested in language evolution sometimes compare "pears and apples", since the behaviours of interest are not correctly (or too broadly) defined. Sometimes there is a tendency to focus on similarities between humans and other primates, while at the same time, differences might be neglected. However, I will point out later why I do not agree with some of the conclusions Thom Scott-Phillips draws from comparative research, with special focus on his chapter four, dedicated to the origins of ostensive communication.

The nature of ostensive signals

Thom Scott-Phillips defines ostensive signals as "signals that express communicative intentions, and hence informative intentions" (p. 9). While the informative intention represents *what* the signaller wants to communicate, the communicative intention conveys the information *that* the signaller wants to communicate. He uses the example of tiling a cup ("I want more coffee") plus establishing eye-contact ("I want to communicate my intention") and then comments that "typically, both types of intention are expressed in one and the same behaviour, such as the tilt of the coffee cup". Based on this example, I am not sure about the nature of ostensive signals: on the one hand, Thom Scott-Phillips

suggests that both intentions are expressed by one behaviour (tilting the cup), but then refers to the importance of eye-contact to signal the communicative intention to a specific recipient and to make clear that one intends to communicate. Without this signal, Thom Scott-Phillips continues, the recipient would not realize that "tilting the cup" was directed at him.

Why am I so picky about this? If we are not clear about what exactly ostensive signals are and which forms they can take in humans (visual gestures? facial expressions? eye gaze? tactile communication? body postures/movements? vocalizations?), then it is also not clear what exactly we are looking for in non-human great apes (or other primates) and how we should interpret findings from comparative research. To give an example: Chimpanzee males use a "penis offer" to communicate that they want to copulate with a female. To make sure that the female perceives this visual behaviour, the males combine it with a "leaf-clipping" behaviour, which causes a sound to attract the female's attention. Thus, while the "penis offer" refers to the *what*, the "leaf-clip" conveys the intention to communicate. One of course could argue that the males only want to change the female's behaviour, but not her mental representations. Furthermore, I am aware that this example only covers the expression, but not the recognition of ostensive-inferential communication (see p. 86), and that particularly in the gestural (in contrast to vocal) modality, we still know very little about great apes' understanding of others' communicative (and informative) intentions (in contrast to vocalizations, as highlighted in the book). Still, I think it is important to point out which types of ostensive signals we would expect to observe in other primates. It is also important to define clearly why we would consider some behaviours, but not others as potential ostensive signals. Gestures, and some vocalizations, are considered, but not facial expressions, which are barely mentioned throughout the book. Facial expressions might not be used intentionally, as are gestures and some vocalizations; however, despite the common notion that facial expressions might merely express internal, affective states and are not intentionally used, this aspect has never been studied systematically by applying the suggested criteria for intentional use to facial communication in non-human primates.

Imagine a chimpanzee approaching another one and using an "arm raise" gesture (ritualized from actual hitting), but combining this gesture with a play face to signal that he wants to play and not attack. The gesture is directed at a specific individual to signal his communicative intent, and the play face informs the recipient about the nature of this approach, since the very same gesture is also used in aggressive contexts. Again, one could argue that it is still unclear whether the signaller is meant to change the other's behaviour instead of mental representations, and that facial expressions are merely expressions of affective states instead of intentionally produced signals. Taken together, what I would like to emphasize is that we have to define clearly which forms ostensive signals can take in other species.

Origins of ostensive communication (Chapter 4)

This particular chapter, from my point of view, does not always represent our current knowledge about primate communication and the corresponding socio-cognitive skills in an appropriate way. While reading it, I got the impression that Thom Scott-Phillips postulated a potential fundamental difference between the communication of humans and other primates, which he then confirms by selecting the corresponding (but not always representative) studies, while neglecting others.

Furthermore, what I missed here was a more in-depth discussion of the function of so-called "attention-getting gestures", which are the tactile and auditory gestures that chimpanzees seem to use to attract the attention of a not-attending recipient. Thus, it could have been of special interest to discuss whether non-human primates use attention-getters to manipulate the attentional states (and thus visual perspectives) of others (conspecifics or humans), thus signalling communicative intentions, and whether signallers combine attention-getters with subsequent specific intention movements, particularly since existing studies draw different conclusions. Furthermore, I think the conclusions drawn from §4.5 on the mindreading skills of non-human primates are rather problematic. The conclusion that "there is little evidence that chimpanzees have command of the type and extent of mental metarepresentations that have been identified as cognitive pre-requisites for ostensive communication" is simply premature. Thom Scott-Phillips even acknowledges that it is difficult to develop a paradigm that is appropriate and ecologically valid for non-human primates and that such studies might be very difficult to conduct because of methodological reasons. I do not want to argue that the socio-cognitive skills of non-human primates are most likely as complex as those of humans and that potential differences are only a matter of quantity, not quality. However – with apologies for this almost trite remark –, "the absence of evidence is not the evidence of absence". In this context it is also not helpful to refer to the possibility that negative findings have been most likely not published. Currently we have to acknowledge that we still know comparably little about the mindreading skills of non-human great apes. The point I want to emphasize here is that we simply don't know (yet) the mindreading skills of non-human primates to the extent they have been studied in humans.

Comments

***

Dan Sperber: From attention-getting to ostension

I found Katja Liebal's comments highly relevant, in part, maybe, because I interpreted them as being more constructive than critical (whether or not they were so intended).

It shouldn't be controversial that humans are capable of indicating to their audience both that they intend to communicate something and what it is that they intend to communicate. Moreover they can do this in such a way that their indicating that they intend to communicate helps their audience recognize what it is that they intend to communicate. If I happen to stare sadly at my

empty glass, my host may notice this and refill my glass, but this is not a case of ostensive communication. If I establish eye contact with my host and then stare sadly at my glass, this is a case of ostensive communication that my host should take as a request for a refill. The two aspects of an act of ostensive communication – informing the audience of one's communicative intention, and informing her of one's informative intention – may be both carried out by the communicative act as a whole (as in most of ordinary verbal communication) or there may be a distinct ostensive cue or signal that informs the audience of the communicator's communicative intention. This is what happens, for instance, in the kind of adult-infant pedagogic interaction studied by Csibra and Gergely. The fact that attracting attention, for instance by establishing eye contact, and conveying a content may be merged or be two distinct aspects of the overall communicative act is a well-known fact and not at all a problem in the study of ostensive communication. This possibility is highly relevant when we look at communication in primates, but it is not something that diminishes the relevance of the ostensive communication framework to ask comparative questions, or it shouldn't be.

Do other animals and in particular other primates also engage in ostensive communication? There is no conclusive evidence that they do but there is at least evidence relevant to asking the question. The case of chimpanzees' penis offer that Katja describes has long been my favourite, with the leaf-clipping looking very much like an ostensive signal and the interpretation of the display of the erect penis, to which the leaf-clipping sound serves to attract attention, as indeed an offer being, favoured by the leaf-clip. Of course, this is not enough to conclude that genuine ostensive communication is taking place: a behavioural explanation without mindreading, let alone, higher order mindreading, is as Katja points out, possible.

Before being able to profitably classify penis offers and comparable cases as fully-fledged/rudimentary/incipient ostensive communication, or as nothing of the sort, it might be more profitable to concentrate on these "attention-getting gestures" Katja mentions. Ostension itself is an elaborate form of attention manipulation and, from an evolutionary point of view, this is where one should look for possible precursors, I believe.

In any case, if the description of typical human communication as ostensive is correct, it is clear that humans are quite unique in using it as much and as richly as they do, whether there is something of the sort in our closest relatives or not. In fact, if there is ostension in other primates, or just if their attention-getting gestures are related to ostension, this would be hugely interesting both from a comparative and evolutionary point of view, and it might help us better understand forms of interaction in humans infants, children, and adults that are not quite ostensive, but close.

Thom Scott-Phillips: Ostension is about mental states, not just behaviour

Katja begins by asking for some clarification about the nature of ostensive signals, so let me try to provide some. Ostensive signals are those that express communicative (and hence informative) intentions. We can, in principle, do this with any behaviour at all, including all those that Katja suggests: gestures, facial expressions, eye gaze, touch, body posture, movement, vocalisations, and indeed anything else. So, to pick up on the example Katja raises, it is possible, in principle, for ostension to be expressed only with a tilt of the coffee cup. However, it is sometimes more appropriate to combine two behaviours in the one signal, and so eye contact is add to the tilt of the coffee cup (this could be because, say, the visual environment is noisy, or because there is a politeness norm that forbids just gesturing to waitresses in an impersonal manner). The two behaviours are produced as two parts of the same signal (which expresses both communicative and informative intentions). It is also possible, as Dan suggests, for two behaviours to serve these two functions independently (i.e., one expresses the communicative intention, the other expresses the informative intention), although I'd suggest that this is relatively rare.

How, then, to interpret the example Katja offers, of a chimpanzee combining a "penis offer" with an attention-grabbing behaviour like leaf-clipping? These two behaviours have functions that are, as Katja notes, superficially similar to the functions of informative and communicative intentions, respectively. The penis offer is the what, and the leaf-clipping is the that. However, this functional similarity does not imply a cognitive similarity. Informative intentions are intentions to change others' representations of the

world. An intention to simply change behaviour is not the same thing, and, as Katja and Dan both point out, it is quite possible that one or both of these behaviours is focused on behaviour rather than mental states. This is why one especially good way to test for an informative intention is to experimentally dissociate the behavioural outcome from the intended change in mental representations (p. 87). I don't know how one might do this in the case of the chimpanzee penis offer, but until somebody does so, and shows that the behaviour is driven by an intention to change mental states, we cannot conclude that the penis offer is an expression of an informative intention. The story for leaf-clipping is similar. If this is an expression of a communicative intention, we need evidence that it is driven by intentions not simply to change behaviour, but rather to encourage others to mentally recognise that the chimpanzee has an informative intention. Again, we don't have this sort of evidence for chimpanzees. If we were to acquire such evidence, I would happily change SOM's conclusion that non-human primate communication is likely not ostensive. This would be true regardless of modality.

Katja concludes her comments with a complaint that SOM's conclusions about chimpanzee mindreading are premature. As she rightly points out, absence of evidence is not the evidence of absence. I fear, however, that there has been a slight over-interpretation here. Nowhere do I claim that it has been shown that chimpanzees do not have command of the type and extent of mental metarepresentations in question. I claim only that "there is little evidence" for such a conclusion. I get the impression that, at present, Katja would agree.

17

COMMUNICATION WITHOUT METAPSYCHOLOGY

By Tad Zawidzki

This is an excellent book. I cannot think of another on this topic that matches its clarity, concision, accessibility, comprehensiveness, and argumentative rigor. I'm quite amazed that Scott-Phillips has managed to combine such seemingly antithetical virtues in one work. The discussion is also admirably honest: Scott-Phillips owns up to the obvious weaknesses with the view and offers strong responses.

I am a little embarrassed and anxious, therefore, because I disagree with most of the main theses of the book. Not all of them. Scott-Phillips persuades me that pure code theories of language origins are hopeless. I am also persuaded that some kind of inference is necessary to explain linguistic communication. There are also persuasive discussions regarding the dearth of combinatorial communication systems in nature, and the role of cultural attractors in the evolution of languages. However, I totally reject the main thesis of the book: that linguistic communication is entirely parasitic on ostensive-inferential communication, where this is understood in terms of metapsychological competence, in particular, the capacity to attribute recursive mental states, via the kinds of inferences that scientists use to infer causal hypotheses from observable data (§1.4).

I detail my reasons for skepticism regarding the metapsychological roots of linguistic communication in the next section. I think that, given Scott-Phillips' background assumption about the only viable alternative theories of language evolution, the case he makes is plausible. However, another problem with the book is the assumption that the code model and Scott-Phillips' version of the ostensive-inferential model exhaust the possibilities. There is a third alternative that has been explored by some philosophers and psychologists: conceiving of language as a shared, normative practice. Scott-Phillips also conflates two questions: what is required for successful linguistic communication, and *how* human beings meet these requirements. He assumes that the requirements on successful communication – the production of relevant signals and their interpretation as such – can be accomplished only via metapsychological inferences to attributions of recursive mental states. But this is not the only possible mechanism for implementing relevance. It appears to be the only possible mechanism if one focuses entirely on mechanisms endogenous to individual communicators, as Scott-Phillips

does. However, when one considers linguistic communication as a shared normative practice, it is possible to specify social mechanisms that obviate the need for metapsychological inferences to attributions of recursive mental states in order to implement relevance. In the last section, I explain this alternative and how it can evade many of the problems I raise for Scott-Phillips' view.

Despite its overall clarity, there is one fundamental topic about which the book could be clearer: what the central thesis is. As near as I can tell, this is the book's central thesis: "the common assumption that the linguistic code makes linguistic communication possible is simply false. Instead, linguistic communication is a type of ostensive-inferential communication, made possible by metapsychology" (p. 21). But what is meant by "makes possible"? I think most would agree, certainly in the wake of Scott-Phillips' arguments against the code model, that linguistic codes are not *sufficient* for linguistic communication. However, even Scott-Phillips must grant that ostensive-inferential communication made possible by metapsychology is also insufficient for *linguistic* communication. Furthermore, if Scott-Phillips' arguments are correct, then surely both are necessary. So the thesis should be that linguistic communication is made possible *both* by linguistic codes *and* by ostensive-inferential communication. This is the view that I will urge below, identifying the different yet equally important roles each plays. I will also argue that ostensive-inferential communication is possible without metapsychology.

How is metapsychology supposed to help?

Scott-Phillips marshals a series of compelling examples illustrating to what extent literal meaning underdetermines speaker meaning. Words, as he repeatedly points out, can be used to mean anything in specific contexts, no matter what their literal meanings are. This is a huge problem for the code model of linguistic communication. But, argues Scott-Phillips, an ostensive-inferential model that puts pragmatics first can avoid this problem. The reason is that interlocutors can infer the specific mental states driving communicative acts in particular contexts, thereby determining precisely what the intended speaker meanings are.

But there are a number of problems with this proposal. First, the book is entirely silent on how a finite, computational system, like the human mind, can successfully infer mental states from observed behaviours and contexts. It is true that literal meaning underdetermines speaker meaning. But it is just as true that observed contexts and behaviours underdetermine the mental responses to and causes of them. This is the well-known problem of holism: any finite observed behaviour or circumstance is compatible with an infinite number of distinct sets of mental states, and any finite set of mental states is compatible with an infinite variety of future behaviours, if we make appropriate adjustments to the other mental states that constitute an agent's whole set. Thus, it is unclear how attempting to infer a speaker's mental states from the circumstances and behaviour that accompany their utterances can help establish what they mean. Or, at least, Scott-Phillips has not sketched a plausible computational mechanism for accomplishing this feat.

The only attempt to formally model propositional attitude attribution of which I am aware concludes that computationally bounded interpreters are not guaranteed to make accurate attributions of propositional attitudes to a target unless they also model the target's reasoning and belief revision strategies (Alechina & Logan, 2010). And perhaps Scott-Phillips is gesturing at this when he likens the process of interpretation to scientific inference. The only problem is that we have no idea how the brain implements such inferences. They are what Fodor calls "isotropic" (1983): any information might be relevant to any inference. And it is hard to model isotropic inferences in a computationally tractable way. This is the so-called "frame problem" of artificial intelligence. I am not saying that this problem is insurmountable, just that Scott-Phillips owes us at least a sketch of how the brain might solve the seemingly intractable problem of interpreting behaviour, if he thinks behavioural interpretation can help resolve underdetermination of speaker meaning by literal meaning.

But there are more serious problems lurking in the background here. Even if we identify computationally tractable algorithms for scientific inference, it is not clear that these are at work in most cases of communication. Scientific inferences are laborious, conscious, time consuming, and often unreliable, or at least epistemically fraught. But, typically, quotidian linguistic interpretation is fast, automatic, highly reliable, and unconscious. And it appears equally effortless with people we know well as with complete strangers who speak the same language. [1] Surely, we have access to more relevant background information about familiar than about unfamiliar people. So inferring communicative intentions should be easier in the former than in the latter case. But, except in special circumstances where we are not using a conventional language, communication seems equally effortless with familiar as with unfamiliar interlocutors. How can this be the case if inferring communicative intentions is like scientific inference, and hence sensitive to the presence/absence of potentially relevant background information? So even if we identify computational mechanisms capable of implementing slow, laborious scientific reasoning, it seems unlikely that routine quotidian communication makes use of these same mechanisms.

One of the main advantages of code models over ostensive-inferential models of communication is that it is relatively straightforward to implement fast, automatic, efficient, and reliable decoding algorithms. As Scott-Phillips points out in an analogy to mathematics, decoding does not seem like a process that requires inference. One might add that it does not seem isotropic: in a code there is only a limited amount of information that could possibly be relevant. This is one way that some computations avoid the frame problem. And it suggests a way in which codes might help overcome the prima facie computational intractability of interpretation. If not just linguistic items, but also the non-linguistic behaviours that accompany them, constitute biological or conventional codes for information relevant to interpretation then this might make the task of interpreting communicative acts more tractable. And this sets the stage for another problem with Scott-Phillips' metapsychological model of ostensive-inferential communication: many of the behaviours to which he refers when explaining how metapsychological inferences help determine speaker meaning are, arguably, themselves codes for emotions, reactive attitudes, and other relevant information.

Raised eyebrows, puffed cheeks, smiles, winks, flared nostrils, raised voices, forceful intonations, etc., can all carry an indefinite variety of meanings on their own. It is only in the context of certain culturally variable expectations that they take on determinate significance. Enculturation is the lifelong process of learning such significances from their repeated demonstration in witnessed conversations and other interactions. Interlocutors from the same culture arrive at complementary interpretations of non-linguistic contexts and behaviours only because they have been shaped by similar cultures to situate them in stereotyped scripts or frames that limit the set of viable interpretations. So such non-linguistic contexts and behaviours come to function as codes that simplify the task of interpretation, obviating the need for metapsychology. We need not engage in metapsychological inferences about what other subway riders expect from us, for example, because we have all been socialized to have complementary expectations in such contexts.

As Olivier Morin has pointed out to me, and as I concede below with respect to eye contact, there is much evidence that some such nonlinguistic communicative behaviours have universal significances across cultures. But it does not follow from this that they require science-like inference to be interpreted: there may also be biological codes linking such behaviours to relevant information. For example, on Csibra and Gergely's "natural pedagogy" hypothesis (2006; 2009; 2011), eye contact, "motherese", contingent interaction, and other such low-level behavioural cues carry a very unambiguous meaning for human infants: they signal the imminent demonstration of generic information regarding an object to which the performer is about to refer. I am not sure why infants require science-like inference to interpret such behaviours in context-sensitive ways; their meaning appears remarkably context-insensitive, as with any code.

The lesson here is the following. If interpreting non-linguistic behaviours, like puffed cheeks, etc., in specific contexts is supposed to help resolve underdetermination of speaker meaning by literal meaning of utterances, then it must do so in a computationally tractable way; otherwise it cannot explain the speed, automaticity, and reliability of typical communication. But a speaker's mental states seem just as underdetermined by such behaviours if all interpreters have to go on is metapsychology: after all, any finite set of behaviours is strictly compatible with any finite set of mental causes, due to holism. This problem can be avoided if non-linguistic behaviours themselves constitute a culturally or biologically determined code for communication-relevant information. But if this is the case, then there is no need for metapsychological inference to help in routine cases of interpretation. Instead, what is needed is mastery of the non-linguistic, embodied communication codes of particular cultures, and of the species as a whole. [2]

The final problem with Scott-Phillips' appeal to metapsychology is one he addresses forthrightly, but in my view, unsuccessfully. He acknowledges that there are apparent counterexamples to the claim that successful ostensive-inferential communication requires metapsychological inference to recursive mental states: preschoolers and autistic individuals. His response is to cite evidence that metapsychological inference might be much less cognitively demanding than people assume. Unfortuna-

tely, the evidence he cites does not establish this. First, the evidence for automatic understanding of indefinitely higher orders of nested propositional attitudes comes from adults, with years of experience interpreting narratives that make explicit such states, using the recursive structures of language. Second, the evidence he cites from implicit false belief tasks with pre-verbal infants does not support the hypothesis that pre-verbal infants understand *recursive* mental states. At best, they show some sensitivity to first-order mental states, like false beliefs, though even this has been questioned, or at least interpreted as a minimal version of the mindreading available to adults (Butterfill & Apperly, 2013). In fact, Scott-Phillips claims affinity between his view and Apperly and Butterfill's (2009) minimal mindreading hypothesis, but they explicitly deny that minimal mindreading makes the attribution of recursive mental states possible. So, Scott-Phillips has given us absolutely no reason to think that pre-verbal infants meet the requirements he claims for ostensive-inferential communication. But clearly they are capable of it.

None of this is an indictment of Scott-Phillips' claim that codes are insufficient for explaining communication, or of the importance of inference to successful communication, or of the relevance-theoretic analysis of communication. It is a critique of his assumption that the inferences that make communication possible must be metapsychological, and that the only way the relevance-theoretic analysis of communication can be implemented is through the attribution of recursive mental states. I now turn to an alternative proposal that eschews metapsychology without neglecting the importance of inference and relevance.

Interpretation without metapsychology

According to the relevance-theoretic analysis of communication, here is what is necessary for linguistic communication on the interpreter's side: "In order to interpret B's utterance, A searches for an interpretation that optimizes relevance i.e. one that maximizes the positive cognitive effects, and minimizes the processing effort required" (p. 59). Here is what is required on the signaller's side: "the Communicative Principle. It states that every ostensive stimulus carries a presumption of its own optimal relevance. What this means is that when signallers produce signals, they produce those signals that maximize the relevance of the stimulus to the audience" (p. 60). Neither of these says anything about metapsychology or the attribution of recursive mental states. Those belong to hypotheses about *how* signallers and interpreters implement relevance, not to a specification of *what* is required for successful communication. And Scott-Phillips elides this distinction. His full description of the requirement that signallers produce optimally relevant signals includes the following: "…they produce those signals that maximize the relevance of the stimulus to the audience, given both the signaller's goals and preferences, and what the signaller knows about the receiver's goals and preferences" (p. 60). But why does the signaller need to know about the receiver's goals and preferences? If there were some other way to produce signals that maximize the relevance of the stimulus to the audience, wouldn't communication be successful, whether signallers knew why it worked or not?

I think Scott-Phillips inherits from both the Gricean and the relevance-theoretic approaches a fixation on atypical and overly intellectualized forms of communication. It is true that there are circumstances in which signallers can maximize relevance only by thinking about what their audience believes about their beliefs and intentions, and audiences can interpret communicative acts only by thinking about what signallers believe about and intend for their beliefs. These usually involve non-standard, jury-rigged communicative signals, especially those targeted at specific audiences in specific circumstances, as when spies need to communicate in ways that others cannot understand. But there is no reason to suppose that this is a good model for most conventional, linguistic communication.

Suppose interlocutors typically think in this way. Conversations are treated as joint activities the goal of which is to share information, and interlocutors are, as a default, simply expected to play the appropriate roles in this joint goal. We have certain evolved behaviours that are highly reliable signals that ensuing behaviours have as their goal the sharing of information that is not public, e.g., eye contact. So, when a signaller wants to share such information, she makes eye contact with her audience. The audience comes to expect that what will follow aims to make manifest information that she does not have and that is optimally relevant to her goals. This is not because she has inferred the mental states of the signaller; rather, eye contact is a highly reliable indicator that this will happen. The signaller then engages in a performance that, given culturally determined assumptions she shares with her audience, ought to make manifest such information to the audience. Again, she does not speculate about what the interlocutor thinks or knows; she simply makes use of a conventional communicative act which, in that context, it would be rational for anyone sharing her cultural background to interpret as making manifest the relevant information. If all goes well, there is no reason to entertain any hypotheses about what the interlocutor was actually thinking. We need only make assumptions about how people ought to respond to certain behaviours in certain circumstances.

It is important here not to conflate relations like being-informed-by and having-a-goal, with mental states like beliefs and intentions. Mental states have traditionally been conceptualized as theoretical posits aimed at causal, quasi-scientific explanation of observable behaviour (Sellars, 1956/1997). To conceive of bodily behaviour as caused by mental states, one must conceive of the body as animated by an enduring, unobservable object of which they are states: the mind. These mental states must be conceived of as interacting in complex, unobservable ways to yield behaviour; otherwise there would be no point in positing them – tracking behavioural patterns would be sufficient. If positing mental states is relevantly like theorizing to unobservable causes in science, then it must support a robust behavioural appearance/mental reality distinction: the hypothesizer of mental states must be capable of conceiving the possibility that qualitatively indistinguishable, counterfactually robust behavioural patterns are caused by *different* mental states, as when medical diagnosticians conceive the possibility that the same patterns of symptoms are caused by different underlying conditions. But there is absolutely no evidence that infant interpreters, or adults in the heat of seamless, dynamic, communicative interaction, are attributing such explanatory, theoretical constructs. All that matters for quo-

tidian interpretation is reliable behavioural anticipation, not limning the true mental causes of behaviour. For these purposes, interpreters need only track what bouts of behaviour are likely informed by, and what they aim at. These are relations between targets of interpretation and non-psychological facts that interpreters can represent independently of any interpretive project. Applying Gergely and Csibra's "teleological stance" (2003), interpreters need only think of behaviours as aiming at some observable alteration in the environment, and as guided by information relative to which the behaviours constitute rational means to that goal, whether the information is actual or not. Interpreters can represent such facts without any concept of unobservable, mental causes of behaviour.

Many human interactions are like this. Consider games like chess. Here, metapsychology is not necessary unless one's opponent starts making really irrational moves. Otherwise, one can simply use the norms that define chess to anticipate one's opponent's moves. Solving crossword puzzles is similar, and it involves linguistic interpretation, like communication. I know nothing about the psychological profiles of the persons who construct the crossword puzzles I solve every morning. What I do know are certain linguistic conventions, like word spelling, and certain culturally specific facts that underlie allusions, puns, and other kinds of cryptic clues. These are enough to infer the correct solution to he puzzle; no metapsychology is necessary. My suggestion is that everyday conversational interpretation is similar to this. We think about what people *ought* to know or infer, and because, due to similar socialization, we largely agree on this, we can communicate without thinking about each other's psychologies.

Of course background knowledge about specific individuals helps (although it is not necessary, as we communicate successfully with complete strangers). But there is no reason to construe even such background knowledge metapsychologically. We can look to recent behaviour, line of sight, manner, personal history, appearance, etc., to infer what a potential interlocutor is or is not likely informed of. But this requires no speculation about mental representations. Such informedness can be conceived of in terms of observable relations, like line of sight, to non-psychological facts of which we are aware, or stereotypes regarding certain types of people and their sensitivity to certain types of facts.

If I am picking berries with someone, it goes without saying (or thinking) that they like berries. If they have not seen the patch from which I am currently picking, or if they do not join me, they are clearly uninformed about the edibility of the berries. I automatically make eye contact, given that we are partners in a joint endeavour that includes the joint goal of sharing information about berries. They automatically expect a behaviour that will make manifest information they do not have that is maximally relevant to their current goals – picking and eating berries, it just so happens. I slowly and exaggeratedly eat the berries I am picking off the bush. They wonder, what could this mean, given that it is maximally relevant to my goals, including picking and eating berries? Ah, they think, those berries are edible! Where is the attribution of recursive mental states?

Of course this episode can be interpreted as involving the attribution of nested beliefs and intentions, as Scott-Phillips does (§3.4), but a far simpler explanation of how relevance is implemented is possible. Suppose the following claims are true of human language users:

(1) They conceive of themselves as obligated to share information relevant to goals their partners in joint endeavours are expected to have.
(2) They can tell through various low-level, behavioural cues whether or not their partners have some relevant bit of information.
(3) They have at their disposal low-level, stereotyped, behavioural signals that, as a matter of fact, whether they think about it or not, indicate to their partners that sharing of relevant information is imminent (like eye contact).
(4) They can follow such cues with performances which, as a matter of fact, whether they think about it or not, are interpreted as and, typically, succeed at making manifest such information

My claim is that the relevance-theoretic requirements on successful information can be met under such circumstances without sophisticated metapsychology. Both biological and cultural evolution can insure that such mechanisms implement relevance without metapsychology. For example, biological evolution yields stereotyped behavioural cues of imminent sharing of relevant information. Cultural evolution yields capacities for context-sensitive performances, including both linguistic and non-linguistic components that, in similarly enculturated individuals, are interpreted as, and typically succeed at making manifest relevant information.

Conclusion

I have much more to say about this highly stimulating and insightful book, but very little room to say it in. For example, I do not think that Scott-Phillips is entirely fair to handicap theories of signalling. There are clearly linguistic phenomena that succumb to this analysis. For example, accents are excellent, costly signals of group membership. They are more costly to produce for people who have not been socialized in a particular linguistic group. In prehistory, analogous forms of communicative "filters" may have been a great way of discriminating between people based on likely trustworthiness or complementary interests. For example, Sosis (2003) proposes that rituals constitute costly signals that can filter reliable cooperation partners from unreliable mimics: mimics will see ritualistic preludes to cooperative endeavours as opportunity costs, while those socialized in a community will see them as routine and hence uncostly. If prehistoric demographics were relevantly analogous to those of contemporary hunter-gatherer societies (Powell, Shennan, & Thomas, 2009; Mellars, 2005; Hill et al., 2011), then individuals likely belonged to nested hierarchies of groups composed of other individuals with whom they had varying degrees of affiliation and familiarity (Caporael, 2001). Besides immediate family members with whom they interacted daily, they also had to cooperate with band-mates, members of hunting teams, and members of larger groups like tribes, with whom they interacted rarely. Despite their rarity, such interactions likely constituted some of the most biologically significant ones:

e.g., mustering war parties or exogamous pair bonding. Complex communicative rituals would have been an especially important form of costly signalling in such contexts.

It is possible that the apparent excess expressive capacity of human language, made possible by recursive grammar, descends from costly rituals used to filter reliable from unreliable group members. This would make these structural aspects of language analogous to birdsong, the structural complexity of which derives from sexual selection for signals of mate quality (Fitch, 2010; Miyagawa, Berwick, & Okanoya, 2013). Such content poor yet structurally complex communication systems can avoid the chicken-and-egg problem identified by Scott-Phillips for code-based models of language evolution (§2.3), since capacities for producing structurally complex calls co-evolve with preferences for them, both in populations that use them to advertise sexual quality, like songbirds, and in populations that use them to advertise for cooperative commitment and competence, like, plausibly, prehistoric human populations. This could explain how humans came to have a communicative code that was structurally complex yet semantically impoverished, a "prosodic protolanguage" as Fitch (2010) calls it. Such a code could have then been employed to make ostensive-inferential (yet, if I am right, not metapsychological) communication properly linguistic.

I also share Olivier Morin's concerns about gossip and reputation as means of stabilizing honest communication: this explanation seems circular. And I do not share Scott-Phillips' skepticism about the significance of understanding shared goals to communication (§3.6), as suggested by what I say above. All of these worries can be traced to what in my view is Scott-Phillips' excessively individualistic orientation. Linguistic communication is seen as a tool that one individual uses to manipulate another, who attempts to insure that the manipulation does not go against her interests. But I think if we conceive of communication as a norm governed practice, evolved through cultural group selection (Henrich, 2004) to improve group coordination via practices of information sharing, then we can avoid some of the problems with Scott-Phillips' focus on sophisticated metapsychology. If people typically have complementary goals, similar assumptions about what is relevant and rational, transparent relations to information, and access to low-level signals of imminent information sharing (like eye contact), then successful communication in the relevance-theoretic sense does not require sophisticated metapsychology. Indeed, it is hard to see how increasing group size and complexity in human evolution can have led to increasingly sophisticated metapsychology, as Scott-Phillips assumes (§6.3), given that it would make the problem of attributing mental states increasingly intractable, as people encountered increasing numbers of completely unfamiliar individuals. It is more likely that our ancestors coped with such demographic changes by instituting normative practices that made group-mates more easily interpretable to each other.

[1] This is not obvious, and it is not obvious how to test this empirically. Also, there are clearly cases where it is false: inside jokes, etc., that only people who know each other intimately understand. But my point is that, typically, linguistic communication among speakers of the same language seems qualitatively similar whether or not it involves interlocutors that know each other well. Think of asking strangers who are native speakers of one's language for directions, or the time of day, or ordering in a restaurant, or countless other quotidian, communicative interactions we take for granted everyday. In my experience, these do not seem different from analogous interactions with people we know intimately. It is no more difficult to ask a complete stranger for the time of day then it is to ask members of one's family.

[2] I thank Olivier Morin for pointing out to me the evidence that many such embodied communicative behaviours have universal, culturally invariable significances.

Comments

***

Thom Scott-Phillips: The limits of the analogy between science and mindreading

I have much more to say about Tad's highly stimulating and insightful comments, but very little time to say it in. Still, this gives me an opportunity to simply cut to two foundational issues that, I think, underlie our various disagreements.

First, Tad questions the priority I grant to ostensive communication, over the linguistic code, in making linguistic communication possible. Second, he believes that mindreading in general, let alone recursive mindreading, is a cognitively laborious activity, unlikely to be involved in normal quotidian communication, and unlikely to emerge early in life (and hence unable to explain young children's competence with ostensive communication).

On the priority of ostensive communication, Tad writes: "I think most would agree…that linguistic codes are not sufficient for linguistic communication. However, even Scott-Phillips must grant that ostensive-inferential communication made possible by metapsychology is also insufficient for linguistic communication… surely both are necessary. So the thesis should be that linguistic communication is made possible both by linguistic codes and by ostensive-inferential communication." This is all true. You cannot have linguistic communication without both ostension and codes. But – and this is important – to develop the conventional codes that are used in linguistic communication you must have ostension first. The only codes that can exist prior to ostensive communication are natural codes, and these are not what we use in language. Moreover, natural codes plus ostension does not equal linguistic communication: it equals the use of grunts, laughs, and other such behaviours in an ostensive way (p. 21). So, yes, both ostension and conventional codes are necessary (by definition), but one is ontologically prior. This is what I meant when I said in my response to Liz's comments that the priority I give to ostensive communication over the linguistic code is, among other things, a conceptual claim.

Second, on mindreading and ostensive communication. Early in *SOM* (p. 12), I used an analogy between mindreading and scientific inference, as a way to illustrate the nature of the problem that ostensive communicators face (i.e., one of inference, rather than deduction). Tad reads this analogy more strongly than I intended. He writes: "Even if we identify computationally tractable algo-

rithms for scientific inference, it is not clear that these are at work in most cases of communication. Scientific inferences are laborious, conscious, time consuming, and often unreliable, or at least epistemically fraught. But, typically, quotidian linguistic interpretation is fast, automatic, highly reliable, and unconscious…even if we identify computational mechanisms capable of implementing slow, laborious scientific reasoning, it seems unlikely that routine quotidian communication makes use of these same mechanisms." I do not think that mindreading is laborious, conscious, time consuming, unreliable. On the contrary, in fact. My analogy was designed to illustrate that mindreaders face the same sort of problem as scientists do (an inferential one), but that does not mean that they solve it in the same way. I think that human mindreading is essentially a perceptual skill: think of it more like vision than like science. The inferential problem that mindreading faces is also faced by vision, and in both cases it is solved, I believe, by functionally-specific computational processes.

Of course, we understand the detail of these processes to a much greater degree in the case of vision than we do mindreading. Tad demands that I supply much more detail here, and I confess that I cannot do this at present. However, as I said in my response to Richard's comments, I was concerned in *SOM* to argue not that we have resolved all the details here, but rather that this view is plausible and, moreover, is the best reading of the present data. On reflection, I could have been more circumspect in expressing this view, but I maintain that this is where parsimony takes us. Debate about the nature of mindreading is an ongoing, cross-disciplinary one. It matters because it is about a question that is fundamental for cognition and culture research: what does human social interaction consists of, and how does it work? Different answers lead to different views about a whole range of topics – including, yes, the evolutionary origins of human communication. I hope that readers can see, at least in outline, how the disagreements that Tad brings attention to stem from this initial point of divergence. Still, neither Tad nor I are alone in our views, and I encourage others to join the conversation.

Tad Zawidzki: Social cognition vs. the visual system

I'm sympathetic to Sterelny's claim that the social domain is very much unlike the visual domain. There are very robust regularities in the visual domain having to do with the behaviour of light, optics, etc. But Sterelny argues that the social domain is highly variable over the course of phylogeny, and across cultures. So it's unlikely that there are the kinds of stable regularities relating behaviour to mental states necessary to evolve a modularized theory.

I might add to that that the visual system tracks observable regularities, not regularities linking unobservable states to observable evidence. But theory of mind involves hypothesizing unobservable states and their relations to observable events. If the visual system employed concepts of photons or electromagnetic radiation, then the analogy might hold. But it doesn't, for good reason – scientific reasoning is not a good model for the quick, automatic computations on which it relies. I actually think there are very robust observable regularities in human behaviour, relating agents to worldly states that constitute the goals of behaviour and what it is informed by. And I think human social cognition can track such observable, though abstract properties of behaviours. But I don't view this as metapsychology, as it doesn't involve attributing unobservable psychological states. And I don't see how you can do recursive mindreading without the latter.

One potential way of experimentally distinguishing between metapsychology and the kind of teleological stance plus normative expectation framework I propose is the following. We should look at how people react to failures of communication (and note differences in such reactions across lifespan). If they're attributing unobservable mental states with causal influence over behaviour, then a failure of communication should immediately lead to a new attempt, informed by a new hypothesis about the interlocutor's mental states. The speaker should show signs of trying to correct an earlier hypothesis about her audience's mental state, and designing a new communicative act, based on this corrected attribution. If they have normative expectations that the communicative act should succeed based on the observable circumstances, then a communicative failure should, perhaps, trigger reactive attitudes, where the speaker somehow faults the audience for not getting it. But I'm just thinking off the top of my head here. In either case, I think that looking at how interlocutors deal with communicative failure might be empirically fruitful.

Olivier Morin: There is no such thing as socialization

Forgive the Thatcherian title – the message below is somewhat less radical. I'd like to draw attention to the weakness of the notion of "socialization" (also called, in the post, "enculturation"), Tad Zawidzki's candidate to replace metapsychology as the key precondition for linguistic communication. I know that Tad has considerably refined that notion (and, in many ways, went beyond it) in his intriguing book, *Mindshaping* (Zawidzki, 2013), but here I'll limit myself to his comment above.

Tad's extremely thorough critique of *SOM*, along with his book, convinced me that the scope and power of "weak" mindreading, without recursion or the attribution of fully fleshed beliefs and desires, is underestimated, and of great promise. He also alerted me to a danger I wasn't aware of: sometimes, when relying on mindreading to solve the indetermination problems solved by communication, we jump out of the frying pan and into the fire, for mindreading is plagued by indeterminacy problems of it all. Exciting ideas that I will be chewing on for a while.

Now, as a student of cultural transmission, what draws me to the study of mindreading and ostensive communication is that, thanks to decades of psychological work, they are relatively well explored. They have conditions of felicity, rather clearly defined. We can pin them down to specific points in time and space. None of this applies to socialization. Allow me to be blunt: socialization is not a proper theoretical term. It is a placeholder concept, a promissory note standing for a theory of cultural transmission that was never truly fleshed out. Four big weaknesses stand out.

No criterion for success

How do we know that a socialization process has succeeded? To answer to such a question would be to list the conditions that an American or a Yanomamö must fulfil to be a good American or a good Yanomamö. Such attempts make us uneasy, for good reasons. Twentieth-century anthropologists, like Marcel Mauss or Margaret Mead, found a ready way around the difficulty: they simply assumed that enculturation always succeeds. That takes care of the problem, but at the cost of weakening the concept. Socialization becomes whatever happens to you when you spend enough time (as a child, but also occasionally as an adult) among humans.

No time limit

Which brings me to a second question: When do we know that one is properly "encultured"? It is sometimes said that we never really know, or that it takes a lifetime (and a village, of course) to be socialised. Tad Zawidzki echoes this view when he calls enculturation "a lifelong process". Does this mean socialization can only be completed with death? I am only half-joking here. The Merina of Madagascar, according to Maurice Bloch, seem to hold exactly this view. Mastering and embodying the values of Merina society truly is a lifelong process, one fully completed only when one goes to join the ancestors in death. To have a good Merina life is, so to speak, to become a dead person. Now, the Merina theory of socialization is more thorough than many others, but it means, obviously, that if socialization is truly a lifelong process, its completion cannot be a precondition for anything we do when we are alive. So, either socialisation is not a lifetime's work (but then we need to know when it ends), or it cannot be what makes human communication possible (at least not for those of us who happen to be alive).

No criterion for individuation

Socialization is a culturalist concept: one is not socialized, full stop, but socialized into culture X or Y. Socialization is like learning a language: it can only be completed inside one particular community, and its benefits can only be enjoyed within it: socializations are plural. Here again, Tad Zawidzki echoes the standard culturalist view. He even says that communication is only possible inside a given culture, not outside it. ("Interlocutors from the same culture arrive at complementary interpretations of non-linguistic contexts and behaviours only because they have been shaped by similar cultures to situate them in stereotyped scripts or frames that limit the set of viable interpretations" – my italics).

There are two well-known problems with this view. The first is, we don't know how to identify or define discrete cultures, a point well put by many anthropologists. To be sure, many people are separated by vast discrepancies of habit, ideology, language, etc. Yet most of this variance cannot be neatly partitioned into distinct cultural blobs. "Cultures" also merge. Globalization is making this obvious for everyone (says a French blogger writing from an international airport to a Polish-American philosopher), but there is no reason to think that cultural distances were neatly partitioned event in the deeper past of our species: as Tad Zawidzki notes, hunter-gatherer social structures are complex, with things like nested hierarchies, exogamous moieties, feuding lineages, and the like. One consequence is that simplistic rules like "acquire your group culture and practice altruism within it" are unhelpful. Who is "my group"? What is "my culture"?

The second problem is, of course, that we can communicate with humans of "different cultures" – and we do. Pidgins, lingua francas, trade languages, Globish, even non-verbal communication between strangers attest to that. I am not denying that cultural misunderstandings can happen, but that is neither here nor there. If Tad Zawidzki's communitarian view of communication were true, cross-cultural communication should not be clumsy, or awkward, or difficult, or recent. It should not be there at all. The usual solution is to assume that any two people engaging in communication share some kind of micro-culture, but in that case "culture" becomes virtually synonymous with "conversation", and "socialization" is just another word for "any kind of interaction". The claim that socialization enables communication becomes utterly trivial. Tad's solution to this is the smart one: he acknowledges the role of non-cultural signifiers, like pointing. This, however, is at odds with the claim that enculturation is necessary for communication. The question then arises: How much does socialisation (as distinct from mere language acquisition) contribute?

No way of differentiating enculturation from regular, protracted interactions

From Tad's account, it would seem that a lifetime of interactions with a spouse or a parent ought to facilitate communication. Isn't it, after all, a kind of enculturation, with shared habits, norms, and even the occasional bits of private languages and reference? (I assume every couple has those, usually too embarrassingly cheesy to reveal.) And yet, Tad Zawidzki explicitly denies that sharing someone's life makes communication any easier. This begs the question, how many people must share a practice till it becomes cultural? And why has cultural shared information the property of making communication work, a property that its non-cultural equivalent lacks? After all, shared information between me and my partner is shared by us – whether or not it is also shared by the rest of our society should not matter when the two of us communicate. Where does mere intimacy end? When would a micro-culture begin?

Again, none of this directly impinges on Tad's central claim, that communication can do without sophisticated mindreading. Yet, if our best alternative to mindreading is enculturation, I put my money on mindreading (broadly construed). Enculturation is a black box. So too is mindreading, but at least the field acknowledges that it is. We have started to open the box. We now know a few things about its workings. We know, for instance, that some of its parts at least cannot be cultural in a strong sense: the detection of goals, for instance, is too widespread in other species, and too precocious for that. Meanwhile, "socialization" serves as a code-name, if not for a blank-slate model of cognition, then for a "wet-sponge" model (to adapt a metaphor from Herder) where people indiscriminately soak up from their surroundings a mysterious fluid called "culture" – a fluid endowed with magical properties that ordinary information does not possess. Let us not return to that model.

Tad Zawidzki: A lot to chew on

Thank you so much for the very enlightening critique, Olivier. Like you, I'm left with a lot to chew on.

Some clarifications: I definitely don't think of enculturation as necessary for communication, for the reasons that you give. The point was about linguistic communication – a kind of communication that is easy, fluid, effortless, automatic, yet still remarkably reliable, employing the same conventions. In my view, there is a difference in kind between what transpires between two travellers who share no language when they communicate, and what transpires between two members of the same linguistic community.

And it seems to me that this difference in kind can be captured rather easily with objective measures. E.g., brain areas engaged during communication, levels of stress, and other physiological markers of cognitive effort, etc.

Of course, Olivier is right that we are simultaneously members of many different cultural groups (though I think this is historically a relatively recent phenomenon). And it's plausible that we evolved a capacity to adapt to different groups as interactions among strangers grew. I've even read of some evidence that prehistoric groups were somewhat short lived, and members often had to assimilate to other groups after theirs disbanded. This would have selected for very efficient and reliable cultural learning. But none of this precludes the scientific study of cultures and socialization. These may be very fluid and vague concepts, but science makes use of those all the time, e.g., the species concept in biology. I think Olivier assumes a false alternative: either there are precise individuation conditions on some category, or it can't be studied scientifically. The history of science completely belies this. I would venture to say that most concepts used successfully in science have ended up being vague. Furthermore, there are terrific candidates for objective measures of culture and enculturation. Here's one: the interaction of social learning processes with critical period effects. It's true that we can learn from our social groups and models throughout our lifetimes. But there are limits to this, apparent in everything from accents, to phonology (e.g., Japanese 'r'–'l' distinction), to, I would argue, basic values and sense of humour. I think it's plausible that this has to do with an interaction between social learning and critical periods (periods of brain development in childhood during which acquiring some trait is particularly easy relative to other periods). This is why it is so hard to acquire an accent after puberty. So one might define a culture or language group in terms of the set of people who have acquired some set of traits (language, accent, basic normative assumptions) during their critical periods. This doesn't mean that others can't acquire these traits, or that such people can't lose these traits; just that it's (relatively) much easier to acquire them inside critical periods than outside, and (relatively) much harder to lose them outside of critical periods than inside.

There are many other such objective measures possible. Just reaction time and error rates in response to basic communicative acts can easily define language groups, or maybe better, communities of communicatively fluid interactants. Some such communities are accessible to adults, kids, or anyone who puts the effort in. Others are open only to people exposed to the appropriate social models at the appropriate times (critical periods). Individuals are members simultaneously of many such groups. One can conceive of human brains as computers running different varieties of "cultural software" for interaction with different kinds of groups. I don't see how this kind of complexity makes notions like enculturation less scientifically tractable. Why does there need to be only one culture one is enculturated into? Why must there be a definite moment at which one counts as enculturated? Of course enculturation can come in degrees. That doesn't mean it's not scientifically tractable or real. And it doesn't mean that there aren't relatively extreme degrees of it that yield close to discrete categories: one either is a native French speaker or not (depending on what social models one was regularly exposed to, during a critical period). Another possible objective measure of culture can be derived from Boyd and Richerson's and Henrich's notion of "prestige bias". When they model cultural evolution, they assume that a basic mechanism of social transmission is made possible by "prestige bias": people imitate those who have the most social status. But judgments of social status are themsleves variable, and based on cultural assumptions about what counts as high vs. low status. In my experience, for example, financial success is a marker of high status and "emulatability" among some groups, and a marker of low status and "non-emulatability" among others. So one possible way of objectively measuring cultural phenomena is via judgments of prestige. Individuals who regard the same social models as high prestige, and worthy of imitation (as measured by automatic dispositions to deference, tendency to attend to them when they speak, preference for interaction with over others, etc.) might be taken to constitute a cultural group.

Perhaps I'm thinking more of the notion of "mindshaping" I defend in my book here, than classical notions of socialization or enculturation, of which I know little. There is a large variety of mechanisms of social learning that appear distinctive of human beings, and that can produce relatively stable groups of communicative interactants among whom mutual interpretation is seamless, efficient, automatic, fluid, and extremely reliable (a "System 1" competence, if you like), and relative to which mutual interpretation among outgroup members is clunky, unreliable, difficult, slow, effortful, conscious. That's all I mean by "enculturation" and even if we haven't yet achieved a consensus definition of it, it takes more than this to convince me that it's not worth trying. If

I may indulge in a bit of autobiography, I'm actually a Polish Canadian, and from childhood I've had to negotiate, on a daily basis, two (what appears to me) radically different cultures. It is very difficult for me to express in words the degree and subtlety of differences in expectations, assumptions, values, etc., that characterized my communicative acts and social interactions with Polish immigrants versus native Canadians growing up. Just the sorts of pragmatic implicatures on which Thom's book focuses: detecting irony, humour, etc., were most challenging. I admit that this may be my own idiosyncratic experience – perhaps I have some sort of mindreading deficit. But, judging from conversations with others in my shoes, and from interactions with other immigrants or children of immigrants whose initial language and culture were that of the old country, it is by no means a rare experience. The degree of stress I experienced over attaining status (as measured by number of friends, ease of interaction, success of attempts at humour, etc.) among native Canadians was considerably higher than among my fellow immigrants and children of immigrants. This was confirmed by others I talked to. These are real, objective properties of mutual interpretability that can be used to define cultural groups, in my view.

18

REFLECTIONS ON THE SPEAKING OUR MINDS BOOK CLUB

By Thom Scott-Phillips

I can say without reservation or qualification that the *Speaking Our Minds* book club was the single most challenging and rewarding intellectual experience of my career to date. Every day for two weeks some very bright and engaged people posted extensive comments on my work, and initiated many excellent conversations. It is a privilege to have one's work be the focus of so much attention and good quality debate, and I would like to express my gratitude to cognitionandculture.net for the opportunity.

So many good points were made that I almost want to rewrite the book! That's an exaggeration, of course, but an online book club really is the ideal way to find out how your readers actually read your work – much better than a published review – and these insights have at the very least provided me with plenty of material for a second edition. I don't think I would change any of the main claims, or the overall structure of the book, but if I were to edit it now, I would elaborate on and clarify many things. I won't here repeat matters of clarification, but this does seem a good opportunity to summarise, in no particular order, some of the most important ways in which I would add to the substance of *SOM* in light of the book club.

Clarify that *SOM* does not claim to solve everything. More than one participant understood *SOM* as making the claim that all the interesting questions one might wish to ask about language evolution are settled and resolved. It was not my intention to suggest this. I do think what *SOM* does is provide the right basic framework within which to frame other issues, but that is not to deny that there are many interesting and important questions that remain unanswered. I regret that this distinction is not explicit.

Show more clearly how pragmatics is neglected in language evolution. I said in the précis that pragmatics is neglected in language evolution, and that one way to read *SOM* is "as a demonstration of just how much we can learn about language evolution by taking pragmatics seriously". Two commentators, Bart and Liz, disputed this, arguing that the importance of pragmatics is already widely

appreciated. I don't agree. It is true that the importance of pragmatics is often acknowledged – but this is, in my view, mostly lip-service. I would work to make this point more graphic in any future edition (unless matters change in the meantime, of course).

Precision on the distinction between code and ostension. The distinction between code model communication and ostensive-inferential communication is central to the thesis of *SOM*. Liz posed a number of challenges for it, and Greg's comments triggered an extended and informative discussion on the details of what this difference actually boils down to. It would be a good addition to *SOM* to describe the difference even more precisely and more specifically than is currently the case.

Further discussion of the communication of non-human primates, and defence of the claim that it is most likely not ostensive. SOM argues that ostensive communication is likely to be uniquely human. Katja and Richard's comments posed a number of questions for this view. As I explained in my responses, I do not think that any of their points or arguments undermine my claims, but I do now see that there are several aspects to the argument that could have been elaborated on, or defended in more depth.

Further discussion of social cognition and communication. In the Preface to *SOM* I explicitly said that I would not much discuss alternative views, in order to retain a focus on the positive case for my own arguments (p. xiv). It is not always easy to get right the balance between criticism and advocacy, and if SOM erred too much on one side or the other, it was on the side of advocacy. This is probably most clearly the case in my discussion of the social cognition involved in ostensive communication, where there is a range of alternative views that I did not much discuss. Tad made arguments in favour of one view in his comments, and Richard touched on this theme too. There is certainly scope for discussion of other views here, and further elaboration and defence of my own. One specific topic I did not discuss at all was the modularity of social cognition, and this in particular would make a good addition.

Some elaboration on cultural attraction. I think that cultural attraction is a critically important idea for cognition and culture studies. Looking back, I don't think *SOM* explains in enough detail just why it's so important, and what it has to offer language evolution. Alberto's comments forced to me to do this, and it would have been good to say more about this in the book itself.

Further discussion of Chomsky, Universal Grammar, and associated issues? The issues around nativism, Universal Grammar, and generativism are so vexed that, to be frank, I was not and am still not even sure how to present them in a way that neither side would object too. Passions run high here, and different views about the value of an evolutionary perspective only add more fuel. The incendiary nature of these debates, and the associated risk of misunderstanding, is one (non-scientific) reason why *SOM* was officially neutral on these matters. Another reason is that I actually suspect that, if and when empirical data resolves these debates, all sides will claim that this is what they were saying all along – but to flesh this intuition out would have required a lot (a lot) more work. Still, in

different ways both Deirdre and Tim prodded at these issues, and thanks to those conversations I can see some ways in which I could have said more, without unnecessarily opening a large can of worms.

The book club conversations also suggested several ideas for future research projects. I'd like to highlight three in particular. In each case, the idea for the project comes not from me, but from the comments of one or more of the book club participants. For these ideas I am very grateful, and I would be delighted to see any of them taken forward (whether by me or by others).

Elaboration of how the first ostensive signals could have emerged. Mathieu posed some excellent questions about how exactly an ostensive communication system can get started, and made some suggestions about how to begin to answer these questions. As I suggested in my reply, elaboration on these points could make for a substantial contribution to the literature.

Elaboration of how reputational effects can come to stabilise a communication system. In *SOM* I argued that reputation keeps human communication stable, but, as Olivier pointed out in his comments, I did not offer any explanation of how such a state of affairs could emerge in the first place. Olivier is right that there are some important coevolutionary questions here. Clark hinted at a similar point. As I said in my response to Olivier, this issue is ripe for a modelling project.

Progress on fundamental conceptual issues? Dan suggested in his comments that the framework I use for defining communication is a promising one for addressing some deep and long-standing conceptual issues about the nature of communication. Relatedly, Ira and Tiffany see potential in my framework for alignment between evolutionary biology and pragmatics about how to think about communication. The pursuit of these possibilities would be a project of substantial philosophical import, of relevance to the foundations of several disciplines.

These two lists together illustrate how the book club touched on all the main themes of *SOM*, and just what a rich and productive intellectual experience it was. I hope that the other participants felt similarly. It was challenging in just the right way, and provided many ideas worth following up. So let me finish by reiterating my thanks to all at ICCI who helped to make this happen, and also to all the other participants. It was a privilege to have one's work subject to so much quality discussion.