

Original Article

The Social Evolution of Language, and the Language of Social Evolution

Thomas C. Scott-Phillips, Language Evolution and Computation Research Unit, University of Edinburgh, UK
EH8 9LL. Email: thom@ling.ed.ac.uk

Abstract: Recent years have witnessed an increased interest in the evolution of the human capacity for language. Such a project is necessarily interdisciplinary. However, that interdisciplinarity brings with it a risk: terms with a technical meaning in their own field are used wrongly or too loosely by those from other backgrounds. Unfortunately, this risk has been realized in the case of language evolution, where many of the terms of social evolution theory (reciprocal altruism, honest signaling, etc.) are incorrectly used in a way that suggests that certain key fundamentals have been misunderstood. In particular the distinction between proximate and ultimate explanations is often lost, with the result that several claims made by those interested in language evolution are epistemically incoherent. However, the correct application of social evolution theory provides simple, clear explanations of why language most likely evolved and how the signals used in language – words – remain cheap yet arbitrary.

Keywords: language, evolution, social evolution, language evolution, communication, honesty

“The potential for semantic confusion is greatest with interdisciplinary research” (West, Griffin, and Gardner, 2007, p. 428)

Introduction

In recent years much has been written about the evolution of the human capacity for language, with significant progress on a number of issues, for example how evolution can help explain language universals (e.g., Kirby and Hurford, 2002), how individuals negotiate upon the meaning of signals (e.g., Galantucci, 2005), and how and why the human vocal tract most likely evolved (e.g., Fitch, 2000). However, the social evolution of human communicative behavior has remained at the stage of speculation, with several hypotheses (e.g., Dessalles, 1998; Dessalles, 2000; Dunbar, 1997; Fitch, 2004; Hurford, 2007; Knight, 1998; Miller, 2000; Pinker, 2003; Ulbaek, 1998) but little data advanced. Worse, much of

this speculation misuses or misunderstands the key concepts and terms of social evolution theory, and as a result a number of the positions taken are incoherent. The purpose of this paper is twofold. First, to highlight a number of specific examples of such misunderstandings; and second, to show that when social evolution theory is correctly applied it produces simple, parsimonious explanations. As I hope to show with a range of examples, there is a real need for such an analysis, for confusion within evolutionary linguistics on matters of social evolution theory is remarkably commonplace.

In order to prepare the ground for the arguments to come, the next section makes explicit a crucial distinction, between *ultimate* and *proximate* explanations (see Mayr, 1961; Tinbergen, 1963). I then turn to the question of why we should talk at all, and show that a number of proposed answers to this question come from a misunderstanding of this fundamental distinction between the ultimate and the proximate. Once this is made clear then the question of why we should talk is answered straight-forwardly and very generally: to influence the (future) behavior of others. I also address the other big question for any social evolutionary analysis of communication; that of honesty. Why are human utterances cheap yet reliable? What keeps the system stable? Contrary to a popular belief within evolutionary linguistics (and in intellectual circles more generally) the handicap principle is not the only candidate answer here. In fact, there are several others, and one in particular – that it is *dishonest* rather than honest signals that carry costs – seems very plausible in the case of natural language. I outline the case for this explanation in the fourth section.

Ultimate and proximate explanations

The difference between ultimate and proximate explanations is a rich source of possible confusion, and so should be clearly delineated at the outset. The difference is typically explained as the difference between the *how* and the *why*: ultimate processes explain *why* certain social behaviors exist, and proximate mechanisms explain *how* that logic is enforced. This explanation risks confusion, however, since for some scholars mechanistic descriptions of behavior are precisely what is sought (this is often the case in Psychology) and are therefore given the status of *why*. The difference is may instead be captured as the difference between *process* and *product*.

Ultimate explanations describe the processes by which the present set of behaviors observed in a population arose and are maintained. It is here that terms like kin selection, reciprocity and the like – the technical terminology of social evolution theory – find meaningful use: as processes that explain the continued presence of behaviors (like altruism) which would be otherwise unexpected under the laws of natural selection. Indeed, such ultimate, process-driven explanations of social behavior are, essentially, the *raison d'être* of social evolution theory. In contrast, proximate explanations describe how the logic of such processes is manifested in particular cases. So if we accept that some behavior of a particular species is the result of, say, kin selection (an ultimate process), then to complete the story we should also ask how the members of that species recognize kin and how they discriminate their behavior accordingly (a proximate product of the ultimate process).

In fact, ultimate and proximate explanations can be further divided: the former includes both *function* and *phylogeny*; the latter *mechanism* and *ontogeny* (Tinbergen, 1952). This four-way distinction, commonly known as the “four whys”, is the methodological foundation of contemporary ethology. The particular distinction that this article draws attention to is that between function and mechanism. The other two whys are

concerned with developmental issues (of the species and of the organism) and are hence not central to the points made here.

All of these explanations are desirable. Indeed, they complement each other. Moreover, because the ultimate and the proximate are different *types* of explanation, they cannot coherently be conflated. Ultimate explanations are not – not *at all* – hypotheses about *how* the logic they describe is manifested. For example, we may state that peacocks have large tails because of the handicap principle (Grafen, 1990; Zahavi, 1975; Zahavi and Zahavi, 1997), whereby an organism pays cost as a guarantee of its honesty: only those peacocks who are of sufficiently high quality can afford the largest, most cumbersome tails. Yet this hypothesis makes no necessary statement *whatsoever* about peacock cognition, except that such cognition must produce behaviors that approximate the underlying logic of the process. In particular, a full representation of the process and notion of a handicap (whatever that might mean in the case of peacocks) is unnecessary. The intentional language (e.g., “The peacock *wants* the peahen to *understand* that he is of high-quality”) that is sometimes used to describe such behavior is, then, only as simple but justified shorthand for the processes described by social evolution theory (Dennett, 1995; Grafen, 1999; Hamilton, 1964). Despite everyday meanings, this language is neutral on whether and how the peacock can entertain desire (c.f. “...wants...”) or comprehension (c.f. “...understand...”). Rather, the terms are used exclusively as an anthropomorphized shorthand to conveniently describe animal behavior. Of course, the logic of the handicap principle must be manifested somehow, and a large part of that story presumably involves the peacock brain, but ultimate explanations (like the handicap principle) can *only* be neutral on this, because ultimate explanations are about process rather than product. The two types of explanation are fundamentally different, and should not be confused or conflated.

However, such confusion does often arise. To illustrate, consider recent work on the evolution of cooperation in humans, and specifically the idea of *strong reciprocity* (Fehr and Fischbacher, 2003; Fehr, Fischbacher, and Gächter, 2002; Gintis, 2000). Strong reciprocity is defined as a combination of “a predisposition to reward others for cooperative, norm-abiding behaviors... [and] a propensity to impose sanctions on others for norm violations” (Fehr and Fischbacher, 2003, p.785). This is a proximate definition, and as such cannot be a solution to the ultimate problem of why such behavior should persist in the face of natural selection (West, Griffin, and Gardner, 2007). Another definition makes it clear that strong reciprocity is a product and not a process: “a strong reciprocator is predisposed to cooperate with others and punish non-cooperators, *even when this behavior cannot be justified in terms of extended kinship or reciprocal altruism*” (Gintis, 2000, p. 169, italics added). Strong reciprocity is therefore, by its own definition, a hypothesis about *how* cooperation might be manifested in humans. It is not, however, a hypothesis about *why* cooperation exists at the ultimate level, for it does not describe any *process* that will maintain the system’s stability. This point is well-appreciated within social evolution circles, but the same is not true of evolutionary linguistics, where strong reciprocity is sometimes invoked to explain the large-scale cooperation that is plainly necessary for language to flourish. Consider the following claim, made by a senior member of the evolutionary linguistics community: “strong reciprocity shows how a weakness in reciprocal altruism can be overcome” (Hurford, 2007, p. 477). Yet any “weakness” in reciprocal altruism must be overcome by some other evolutionary process, not by the

product of that process. It is not enough to say that strong reciprocity will do the work where reciprocity alone is insufficient, because such an explanation says nothing about why such a set-up would be stable – which is precisely the problem that reciprocity is sought to explain. The conflation of these two complimentary but different explanations is epistemically incoherent.

This is not the only example, by any means. The next section surveys some of the various hypotheses that have been proposed in response to the question of why humans should talk to each other at all, and shows that conflation of proximate and ultimate explanations is remarkably commonplace. However, when we keep the distinction clear then a simple and very general answer presents itself: we talk to influence the future behavior of others.

Why talk?

When we talk we pass information onto our conspecifics, some of which will be of value and use to them. Yet talking carries production costs; if nothing else there is the opportunity cost of the time spent in conversation (Maynard Smith and Harper, 1995). Why, then, do we talk? Of course, communication most obviously allows us to coordinate our behaviors with far more ease than otherwise, but our communicative instincts go far beyond this. We have a “robust and passionate urge of some kind to communicate” (Bates, 1994, p. 139) with the result that a surprisingly large proportion of the time we spend in conversation is spent on those relatively mundane and trivial matters that we call gossip (Dunbar, Duncan, and Marriott, 1997). Why, then, are we such enthusiastic users of this skill? This question may sound naive at first blush, but no evolutionary answer is immediately forthcoming. Indeed, any real naivety here lies in “linguistic communism” (Bourdieu, 1991, p. 43): the assumption that humans will simply talk to each other for no reason other than the altruistic sharing of information. Correspondingly, several researchers have correctly identified the matter of why we talk at all as a key question for evolutionary linguistics (Dessalles, 1998; Hurford, 2007; Knight, 1998).

The well-known processes of kin selection and/or reciprocal altruism have both been proposed as solutions, or partial-solutions, to this problem. (Reciprocal altruism is in fact a misnamed theory, for it is not a type of altruism at all. On the contrary, what the theory purports to explain is why superficially altruistic acts are in fact *not* altruistic. A better term is *reciprocity* (West, Griffin, and Gardner, 2007), and I will use this hereafter). On this line, large-scale communication is the result of a process whereby we share information either with kin or with those whom we expect to return the favor. With regard to natural language, the problem with the kin selection hypothesis is plain, and acknowledged by the most complete argument presented in its favor: “the glaringly obvious flaw in this hypothesis is that today, language is not used exclusively or even predominantly to communicate among kin” (Fitch, 2004, p.292). Reciprocity is a more popular explanation, endorsed by several researchers (e.g., Hurford, 2007; Knight, 1998; Pinker, 2003; Ulbaek, 1998): “only under the extraordinary conditions of reciprocal altruism can information-sharing take place without loss of fitness to the speaker” (Ulbaek, 1998, p. 41). In order to evaluate this claim, let us be clear exactly what it states: that our willingness to talk is the result of, and is maintained by, a process whereby we *exchange* conversation on a reciprocal basis. However, our conversational behavior does not support this conclusion: if the donation of information were costly and the receipt of information

beneficial, as the reciprocity hypothesis supposes, then the cheating strategy – the defect option in game-theoretic terms – would be to absorb the information offered by others and not to speak at all (Dessalles, 1998, 2007). Yet such a strategy does not trigger anything that looks like cheater detection: those that speak only occasionally are not excluded from future conversations (on the contrary, those most likely to be excluded are those that talk a great deal but do little in the way of listening, yet if speaking were altruistic then these would be generous individuals indeed). The reciprocation argument depends upon the exchange of acts in which one party – supposedly the speaker in this case – pays a cost in order that the other party – the listener – receives a payoff. Yet if cheater detection is any guide to the existence of such relationships then this does not occur in the case of conversation.

Moreover, there are further reasons to suppose that speaking has benefits for the speaker, even when unreciprocated. One is our anatomical divergence from other primates: our vocal tract is significantly different to both our primate cousins and our ancestors, mainly due to the descended larynx, while our ears are not. Although it has been argued that the descended larynx is an adaptation for speech (e.g., Lieberman, 1984), the most compelling explanation is that it evolved to exaggerate size, as it has in other species: a descended larynx produces greater formant dispersion, which is negatively correlated with size (Fitch, 1999, 2002). Vocalizations thus produce some payoff for the speaker, at least in some contexts. As for ears, if information transfer were the function of language, then we could reasonably expect our ears to have developed adaptations to better capture all the input that they can, so that background chatter was no longer just noise but a set of many distinct information sources (Miller, 2000). Yet this is not the case. Of course, these thoughts are pure speculation, but the reasoning is certainly not fanciful. At the very least, they are questions that must be addressed by those who suppose that talk is simply the provision of useful information.

A computational model poses a similar question: if the payoff to communication were simply that the listener gains information, then we should predict more true synonymy and less homonymy (Hurford, 2003). The model works as follows. Agents send and receive signals from one another and are evolved according to one of two possible selection pressures: their fitness is defined either as their ability to signal their intended meaning correctly, or as their ability to deduce intended meaning correctly. Under the first condition the resulting language is free of true synonymy and rich in homonymy, just as we observe in all natural languages. Under the second condition, however, the resulting language was rich in synonymy and free of homonymy, a situation never observed among natural languages. The emergent languages hence resemble real languages only when selection is placed on the agent's productive capabilities, and not their receptive capabilities. This suggests, like the anatomic observations above, that speaking is not the altruistic act of information sharing that a naïve reading would suppose it to be. Instead, the speaker is likely to have gained something from the interaction, just as the listener will have done. It is therefore hard to conclude that the stability of speaking is dependent on reciprocation, a process by which individual speech acts are costly but are maintained through a process of exchange. On the contrary, they suggest that the speaker must receive some net payoff from each speech act. In light of these arguments, and the lack of cheater detection discussed above, reciprocation seems a highly unlikely explanation of our enthusiasm to converse (Scott-Phillips, 2006).

One scholar similarly observes that “a human speech-community is not a personal mutual aid network but is typically an extended group transcending the limits of affiliation on the basis of residence, economic co-operation or kinship” (Knight, 1998, p. 75). However, this observation is then followed by just the sort of error discussed earlier: conflation of the proximate and the ultimate. Specifically, it is suggested that the problem was overcome by a “profound coalitionary restructuring [of human sociality]”. This is a proximate explanation – the structure of human sociality is a *product* of natural selection, not a *process* – and therefore it cannot answer the ultimate problem of why we should talk to each other at all.

This is not an isolated example. Consider the suggestion that our syntactic abilities are based upon the same cognitive mechanisms that must underpin reciprocal altruism (Bickerton, 1998). The assumption here is that because reciprocity is about who did what to whom and when, then there must be some cognitive representation of the corresponding social calculus, and it is this representation that it is suggested was exapted for use in syntax: “thematic analysis *must* have been selected-for long before the hominid-pongid split” (Bickerton, 1998, p. 351, italics added). This is akin to saying that because ant behavior can be explained in terms of kin selection then ants *must* have some of the cognitive machinery required to do arithmetic (see also Dawkins, 1979). In fact, nothing of the sort is entailed by these processes. (For the record, ants appear to recognize kin through smell (Keller and Ross, 1998)). To repeat the previous section’s leitmotif: social evolution theories (kin selection, reciprocity, etc.) are utterly silent – *can only be silent*, by their very nature – on how the logic they describe is actually enacted in any given example. Social evolution theory may of course guide our intuitions in this regard, but no more than that. Although it may indeed be the case some general social reasoning skills provided a pre-adaptation for syntax (c.f. Seyfarth, Cheney, and Bergman, 2005) this is not a necessary condition; it does not follow as a matter of course. Any such claim must do considerably more than simply observe similarities in the structures of our theories of syntax and social behavior.

If it is hard to envisage how reciprocity could occur without a representation of who did what to whom and when, then ask the same question about, say, incest avoidance. Our brains are not pre-wired to represent our family-tree. The evidence instead suggests that a simple heuristic, namely “Do not have sex with anybody who grew up in the same household as you”, does the job to a high-enough degree of approximation (Shepher, 1971, 1983; Wolf, 1970, 1995). Similarly, if humans do engage in reciprocity, there is no reason to suppose that it will not be based upon some similar heuristic (possibly, ironically enough, one that makes use of linguistic markers like, say, accent (Nettle and Dunbar, 1997)). It is not the case that “*of course*, reciprocal altruism theory *relies upon* animals having a capacity to remember, in detail, the past behavior of others” (Hurford, 2007, p. 425, italics added).

This conflation of the proximate and the ultimate, which, as observed in the previous section, renders such claims technically incoherent, is not the only error that evolutionary linguists have made with social evolution theory in general and reciprocity in particular. Consider the following: “If we inform only those people who are likely to return the favor, both of us can gain the benefits of trade...” (Pinker, 2003, p. 28). This is the reciprocity hypothesis (c.f. “...benefits of trade...”). This is then followed by “...it seems clear [though no data is offered – TSP] that we do use our faculties of social cognition to

ration our conversation to those with whom we have established a non-exploitative relationship” (Pinker, 2003, p. 28). This is a more general statement, and not the same thing as the reciprocity hypothesis, which specifically states, when it is used correctly, that we *exchange* pieces of information with each other. One is tempted to conclude that at least some of the popularity of the reciprocity idea is due to an imprecise, overly general use of the term.

That is not to deny that the establishment of equitable relationships was an important landmark in the phylogeny of the language faculty. It is just that such relationships are unlikely to have arisen within the tight and specific constraints of reciprocity. A more general explanation is more likely, and in the next section I will argue for a major role for reputation. Before then, however, I wish to briefly comment on two of the selfish suggestions as to why we talk without relying on reciprocation, both drawn from the evolutionary psychology literature. Doing so will enable us to see more clearly that the “incredibly boring” (Pinker and Bloom, 1990, p. 708) conclusion that language evolved under a pressure to communicate complex propositions is most likely the correct one.

First is the thesis that gossip acts as a form of social glue in the same way that grooming does in non-human primates (Dunbar, 1993, 1997). Although this certainly has a degree of plausibility, two comments should be made. The first is that gossip/grooming is by no means all that we use language for. We frequently talk to others even without any desire to build or maintain social relationships. Indeed, we occasionally talk to people we would like to break ties with. Second, and more pertinently, this hypothesis does not explain why language need be as complex as it is: if social grooming was the sole function of language then grunts could perform the same task, but instead we have a far more cognitively complex and evolutionarily costly communicative mechanism (Pinker, 2003).

A second prominent idea is sexual selection. The suggestion that this process played a prominent role in the evolution of language goes back as far as Darwin (1871) but it, too, cannot tell the whole story, a fact acknowledged by its most prominent modern-day advocate (Miller, 2000). After all, we use communication for far more than attracting mates. That does not rule it out completely, of course. Neither do the following two arguments, both of which have recently been argued to speak against the sexual selection hypothesis (Fitch, 2004): first, that sexually selected traits are usually sexually dimorphic, with the displaying sex (typically, but not necessarily, the male) expressing such traits exclusively, or at least to a far greater degree, than the choosing sex (typically, but not necessarily, the female); second, that sexually selected traits are not usually expressed until puberty.

With regard to the first point, the degree to which we should expect to witness sexual dimorphism is in large part determined by the mating system employed by the species, and in particular the relative levels of parental investment in child rearing: the more investment a parent has to make, the less likely they are to display sexually selected traits (Andersson, 1994; Cartwright, 2000). In humans the norm includes a significant role for *paternal* as well as *maternal* investment in the young. We should therefore expect both sexes to display sexually selected traits. With regards to the second point, it may in fact be argued that full linguistic competence does indeed not develop until puberty: language is, in its widest sense, more than the ability to produce coherent grammatical utterances; it is also the ability to use such utterances in socially appropriate ways, a skill that does not reach adult levels of proficiency until the end of adolescence. Thus, this apparent criticism

becomes a strength of the sexual selection hypothesis; considered in its widest sense, language does indeed develop fully only when sexual selection theory predicts that it will.

In conclusion, then, we may only be equivocal: sexual advertisement may, like gossip, have a role to play. Yet still, these two possibilities (and any possible effects of kin selection) do not at all exhaust the range of domains in which we speak. In short, we put language to an indefinitely large number of uses. This is true of many human traits: “hands may be very good for throwing rocks, shooting arrows, and pressing triggers, but we not infer that manipulability therefore arose as an adaptation for success in aggression” (Gould, 1991, p.53). Indeed; instead we simply conclude that such manipulability is an adaptive end in its own right; a versatile tool with a wide range of applications. Similarly, we use a range of linguistic devices – imperatives, declaratives, requests and the like – to fulfill a diverse range of sociobiological functions, amongst them sexual advertisement and grooming. What these functions have in common is that all involve the alteration of another’s behavior. Rather than claim that any single one is *the* function of language, we should instead draw the general and banal conclusion that language evolved for the purposes of communication (c.f. Pinker and Bloom, 1990); after all, the functional alteration of another organism’s behavior is precisely what it means to communicate (Maynard Smith and Harper, 2003). In sum, language’s great expressivity reflects its diverse selection pressures.

Does such a conclusion actually answer our initial question? We talk to influence the behavior of others, but if that is the whole story then why should others pay attention to our utterances? Indeed, what stops us from simply making claims without any substance, simply so that others will believe them and act accordingly? In the conventional terms of social evolution theory, how is honesty maintained in such a system? That question is the focus of the next section.

Honesty

The theories discussed in the previous section – kin selection, reciprocal altruism, etc. – are those used by social evolutionists to discuss animal social behavior in its many forms, including communication. But communication also brings its own, unique problem, and with it its own set of theories. That problem – honesty, or reliability – is arguably *the* central problem for any evolved signaling system: if there is no mechanism by which individual signals are kept honest then free-riders can invade. If one can gain through the use of a dishonest signal without paying additional costs then we should expect natural selection to favor such behavior. Consequently, signals will cease to be of value, since receivers have no guarantee of veracity. Ultimately, dishonesty will produce listeners who will not attend to signals and the system will collapse in an evolutionary retelling of Aesop’s fable of the boy who cried wolf.

Before we consider this matter for the case of human communication, a brief terminological aside is merited. In its everyday use, *honesty* makes reference to the relationship between a proposition and its truth value. This is roughly the meaning used in social evolution theory, with the obvious but important caveat that here its use is necessarily metaphorical; it does not suppose that an animal must have “meanings” that are either true or false. Rather, terms like honesty are, like the intentional verbs used to describe animal social behavior more generally, a convenient shorthand that makes use of an anthropomorphic gloss of the animal’s behavior. We assign an “intended” “meaning” to the behavior and this allows us to subject it to evolutionary analysis, but this does not at all

suppose that the animal necessarily has “intentions” or “meanings” in any psychologically real sense. In many ways, this is the same point as the ultimate/proximate distinction already discussed above; honesty is used within social evolution theory in an ultimate and not proximate sense.

We may now ask what keeps human communication stable. That is, what prevents dishonesty becoming so widespread that trust in others’ utterances simply breaks down? This question is surprisingly little addressed by language evolution researchers, mainly because those who have thought about language in the context of social evolution theory have, as discussed in the previous section, mainly argued that linguistic use can be explained through reciprocity or kin selection, processes that address the problem of honesty before it is asked. But if we accept that these are unlikely explanations of our willingness to talk then we must look elsewhere for a Darwinian explanation of honesty. Outside of social evolution theory it is sometimes thought that the handicap principle (Grafen, 1990; Zahavi, 1975; Zahavi and Zahavi, 1997) is the only process by which cheap signals may remain honest, but this is not the case: cost-free honesty does exist in nature, and not just in humans. Many male passerines, for example sparrows, typically display dominance badges on their plumage; the larger the badge, the greater the bird’s RHP (Resource Holding Potential; an indexical measure of all factors that influence fighting ability (Parker, 1974)). However, there appears to be no cost associated with the badge, and no obvious barrier to falsification (Rohwer, 1975; Whitfield, 1987).

What, then, are the alternative processes that may explain such a phenomenon? Several have been identified by social evolution theorists (Maynard Smith and Harper, 2003):

- *Indices* – signals that cannot be faked, because the form of the signal is causally related to the signal’s meaning.
- *Coordination games* – where each party has a different preference for the outcome of the interaction, but where an overriding common interest is shared (Maynard Smith, 1994).
- *Repeated interactions, or reputation* – where individuals meet each other repeatedly over time, it may be in both parties’ longer term interests to be honest rather than take whatever short-term payoff may be available through dishonesty (Silk, Kaldor, and Boyd, 2000). Depending upon the specifics of the relationship, the most optimal strategy may be generally honest with occasional deception (Axelrod, 1995; Axelrod and Hamilton, 1981).
- *Punishment of false signals* – if dishonesty is punished then this obviously reduces or nullifies any possible benefit (Clutton-Brock and Parker, 1995).

Of these, it is clear that indices cannot offer a solution to our question, since linguistic form is unrelated to meaning. The handicap principle seems similarly unlikely: utterances are metabolically cheap. What of the other three suggestions? A first observation is that they are not strictly independent of one another. A loss of reputation, for example, whilst ostensibly coming under the banner of repeated interactions, could also be seen as a punishment. In fact, something very similar has been proposed as a solution to the problem at hand.

The crucial idea behind this proposal is that costs are paid not by honest individuals as a guarantee of their honesty (*à la* the handicap principle) but by dishonest individuals as a punishment for their dishonesty (Lachmann, Számadó, and Bergstrom, 2001). In this

scenario, it is cheap and easy to deceive, but there are costs to be paid for doing so. In game-theoretic terms, the costs are paid *away* from the equilibrium; they are paid by those who deviate from the evolutionarily stable strategy (ESS). This contrasts with costly signaling, in which the costs are paid *at* equilibrium; costs are paid as part of the ESS, and the system works because those costs are correlated with that honesty. Consider the paradigmatic example of costly signaling, the peacock's tail. Here, costs are paid by each peacock; the greater the costs, the greater the guarantee, and honesty is guaranteed by those very costs.

This is different to the dominance badges of sparrows because sparrows will pay costs *away from* the equilibrium; in other words, if they *deviate* from honesty. The sparrow's signal is one about RHP, and this is easily verified by other sparrows; they may simply approach more closely or, at worst, initiate a fight. In this way, sparrows are deterred from dishonest signaling; such behavior will result either in fights they are unlikely to win (if they signal a higher RHP than is true), or in less sexual opportunities (if they signal a lower RHP than is true). In the peacock example, however, the signal is one about genetic quality, and the peahen cannot confirm this until they have sired and raised the peacock's offspring. By this time, any sanctions the peahen may impose will be trivial in comparison to the fitness payoff received by the peacock. Hence, the peacock must pay some cost as a guarantee.

In other words, if signals can be verified with relative ease, and if there are costs attached to being revealed as a liar, then cost-free signaling can be evolutionarily stable. In situations where these conditions do not hold, we should expect signals to be costly. Specifically, this will apply in interactions where the protagonists are unlikely to meet again or when communicating about otherwise unverifiable information (Lachmann, Számadó, and Bergstrom, 2001). Of course, dishonesty may still be the more appropriate strategy if the expected payoff is sufficiently large that it outweighs the expected costs. In other words, deception may still pay under certain circumstances, specifically when the benefits associated with doing so outweigh the social costs that will follow.

Hence, honesty can be ensured if dishonesty is costly. One important implication of this analysis is that it allows for signals to take an arbitrary form (*ibid.*). It has previously been assumed by some language evolution researchers that signals cannot be arbitrary because they are open to invasion from cheats: "resistance to deception has *always* selected against conventional [arbitrary – TSP] signals – with the one puzzling exception of humans" (Knight, 1998, p. 72, italics added). However, if the costs associated with a signal are removed from the signal form – that is, when the costs are paid by the *dishonest* individual – then the requirement that signal form is causally related to meaning is removed with it. Consider again the sparrow's badge of status; not only is it honest but it is also arbitrary. Given the already discussed implausibility of other mechanisms that have been proposed to explain the emergence of arbitrary symbolism, it is tempting to suggest that the imposition of some form of cost for dishonest signaling is a necessary (but obviously not sufficient) condition for such a process to take place.

I know of no attempts to test this idea with human subjects. In fact, evolutionary linguistics has thus far shown little interest in Lachmann, Számadó and Bergstrom's work, despite its evident power and plausibility. This is not due to a lack of exposure. The work was published in a major journal (*Proceedings of the National Academy of Sciences*), and is cited extensively elsewhere (41 times as of June 2007). However, it is not cited at all in the

collections of work that have arisen from the biannual conferences on the evolution of language that have taken place since the article's publication (Cangelosi, Smith, and Smith, 2006; Tallerman, 2005; Wray, 2002), nor in an edited book whose explicit goal was to "bring together... all the major perspectives on language evolution...", with each chapter written by "key authorities..." so that "...together they cast the brightest light yet on questions surrounding the origin and evolution of language" (Christiansen and Kirby, 2003, p. 2). Despite these goals the problem of what maintains honesty is barely addressed and where it is the most plausible and likely solution is not discussed. One inevitably wonders if this neglect is simply because the explanatory power of the suggestion, and of social evolution theory more generally, is yet to be realized by many language evolution researchers.

Conclusions

I have sought to make three crucial points. First, I have made clear the distinction between ultimate and proximate explanations of behavior: the first is concerned with the process of evolution, the second with its product. These are different types of explanation that, when used correctly, will support and reinforce each other. However, the two are often conflated within the language evolution literature, a practice which leads to epistemically incoherent claims.

Second, even when understood correctly, social evolution theory is often over-applied in evolutionary linguistics. A specific example is reciprocity, which in many cases appears to be used in a fuzzy way to refer to general social cognition. However, a brief analysis of linguistic behavior shows that its specific claims – that speaking is costly when unreciprocated – are highly unlikely to be true.

This conclusion opens up the question of honesty. Contrary to a common belief outside social evolution, the handicap principle is not the only explanatory candidate here. My third central point is therefore that one specific alternative – that individuals pay costs for *dishonesty* – seems very plausible in the case of humans. Bafflingly, this suggestion has not yet been taken up by the evolutionary linguistics community, despite significant exposure elsewhere. The need to find an evolutionary explanation for honesty is doubly important in the case of humans, since multilevel selection means that our individual fitness will be affected not only by our own honesty but by the overall level of honesty in our social group. The individuals of any group polluted by a high degree of dishonesty – or any other socially divisive behavior – will lose faith in the notional social contract. This break-up of socially-mediated cooperation will most likely lead to selection against the group in question, in addition to the fitness costs paid by the individual members.

It is hard to escape the conclusion that social evolution theory has been too often misused by language evolution researchers. This is regrettable, for when applied correctly it is a powerful explanatory tool. It is hoped that this article has shown how that power might be correctly applied to language.

Acknowledgements: I wish to thank Jim Hurford, Andy Gardner, Andrew Smith, Stu West, Tamsin Saxton and two anonymous reviewers for valuable comments and discussion on the issues raised. I am funded by award number 117604 from the Arts and Humanities Research Council of Great Britain.

Received 27 June 2007; Revision submitted 1 October 2007; Accepted 8 October 2007

References

- Andersson, M. B. (1994). *Sexual Selection*. Princeton, NJ: Princeton University Press.
- Axelrod, R. (1995). *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, R., and Hamilton, W.D. (1981). The evolution of cooperation. *Science*, 211, 1390-1396.
- Bates, E. (1994). Modularity, domain specificity and the development of language. *Discussions in Neuroscience*, 10, 135-156.
- Bickerton, D. (1998). Catastrophic evolution: The case for a single step from protolanguage to full human language. In J.R. Hurford, M. Studdert-Kennedy and C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 341-358). Cambridge: Cambridge University Press.
- Bourdieu, P. (1991). *Language and Symbolic Power*. Cambridge: Basil Blackwell.
- Cangelosi, A., Smith, K., and Smith, A.D.M. (Eds.). (2006). *The Evolution of Language*. Singapore: World Scientific Publishing Company.
- Cartwright, J. (2000). *Evolution and Human Behavior*. New York: Palgrave.
- Christiansen, M.H., and Kirby, S. (Eds.). (2003). *Language Evolution*. Oxford: Oxford University Press.
- Clutton-Brock, T.H., and Parker, G.A. (1995). Punishment in animal societies. *Nature*, 373, 209-216.
- Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. London: John Murray.
- Dawkins, R. (1979). Twelve misunderstandings of kin selection. *Zeitschrift für Tierpsychologie*, 51, 184-200.
- Dennett, D.C. (1995). *Darwin's Dangerous Idea*. London: Penguin.
- Dessalles, J.L. (1998). Altruism, status and the origin of relevance. In J.R. Hurford, M. Studdert-Kennedy and C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 130-147). Cambridge: Cambridge University Press.
- Dessalles, J.L. (2000). Language and hominid politics. In C. Knight, M. Studdert-Kennedy and J.R. Hurford (Eds.), *The Evolutionary Emergence of Language* (pp. 62-80). Cambridge: Cambridge University Press.
- Dessalles, J.L. (2007). *Why we Talk: The Evolutionary Origins of Language*. Oxford: Oxford University Press.
- Dunbar, R.I.M. (1993). Co-evolution of neocortex size, group size and language in humans. *Behavioral and Brain Sciences*, 16, 681-735.
- Dunbar, R.I.M. (1997). *Grooming, Gossip, and the Evolution of Language*. London: Faber.
- Dunbar, R.I.M., Duncan, N.D.C., and Marriott, A. (1997). Human conversational behaviour. *Human Nature*, 8, 231-246.
- Fehr, E., and Fischbacher, U. (2003). The nature of human altruism. *Nature*, 425, 785-791.
- Fehr, E., Fischbacher, U., and Gächter, S. (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature*, 13, 1-25.
- Fitch, W.T. (1999). Acoustic exaggeration of size in birds by tracheal elongation: Comparative and theoretical analyses. *Journal of Zoology (London)*, 248, 31-49.

- Fitch, W.T. (2000). The evolution of speech: A comparative review. *Trends in Cognitive Science*, 4, 258-267.
- Fitch, W.T. (2002). Comparative vocal production and the evolution of speech: Reinterpreting the descent of the larynx. In A. Wray (Ed.), *The Transition to Language* (pp. 21-45). Oxford: Oxford University Press.
- Fitch, W.T. (2004). Evolving honest communication systems: Kin selection and mother tongues. In D.K. Oller and U. Griebel (Eds.), *The Evolution of Communication Systems: A Comparative Approach* (pp. 275-296). Cambridge, MA: MIT Press.
- Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive science*, 29, 737-767.
- Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, 206, 169-179.
- Gould, S.J. (1991). Exaptation: A crucial tool for an evolutionary psychology. *Journal of Social Issues*, 47, 43-65.
- Grafen, A. (1990). Biological signals as handicaps. *Journal of Theoretical Biology*, 144, 517-546.
- Grafen, A. (1999). Formal Darwinism, the individual-as-maximizing-agent analogy and bet-hedging. *Proceedings of the Royal Society of London, series B*, 266, 799-803.
- Hamilton, W.D. (1964). The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7, 1-52.
- Hurford, J.R. (2003). Why synonymy is rare: Fitness is in the speaker. In W. Banzhaf, T. Christaller, P. Dittrich, J.T. Kim and J. Ziegler (Eds.), *Advances in artificial life - Proceedings of the 7th European Conference on Artificial Life (ECAL), lecture notes in artificial intelligence, Vol. 2801* (pp. 442-451). Berlin: Springer Verlag.
- Hurford, J.R. (2007). *Origins of Meaning*. Oxford: Oxford University Press.
- Keller, L., and Ross, K.G. (1998). Selfish genes: A green beard in the red ant. *Nature*, 394, 573-575.
- Kirby, S., and Hurford, J. R. (2002). The emergence of linguistic structure: an overview of the Iterated Learning Model. In A. Cangelosi and D. Parisi (Eds.), *Simulating the Evolution of Language* (pp. 121-148). London: Springer Verlag.
- Knight, C. (1998). Ritual/speech coevolution: A solution to the problem of deception. In J.R. Hurford, M. Studdert-Kennedy and C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 68-91). Cambridge: Cambridge University Press.
- Lachmann, M., Számadó, S., and Bergstrom, C.T. (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences*, 98, 13189-13194.
- Lieberman, P. (1984). *The Biology and Evolution of Language*. Cambridge, Mass.: Harvard University Press.
- Maynard Smith, J. (1994). Must reliable signals always be costly? *Animal Behaviour*, 47, 1115-1120.
- Maynard Smith, J., and Harper, D.G.C. (1995). Animal signals: Models and terminology. *Journal of Theoretical Biology*, 177, 305-311.
- Maynard Smith, J., and Harper, D.G.C. (2003). *Animal Signals*. Oxford: Oxford University Press.
- Mayr, E. (1961). Cause and effect in biology. *Science*, 134, 1501-1506.
- Miller, G.F. (2000). *The Mating Mind*. London: BCA.

- Nettle, D., and Dunbar, R.I.M. (1997). Social markers and the evolution of reciprocal altruism. *Current Anthropology*, 38, 93-99.
- Parker, G.A. (1974). Assessment strategy and the evolution of animal conflicts. *Journal of Theoretical Biology*, 47, 223-243.
- Pinker, S. (2003). Language as an adaptation to the cognitive niche. In M.H. Christiansen and S. Kirby (Eds.), *Language Evolution* (pp. 16-37). Oxford: Oxford University Press.
- Pinker, S., and Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, 13, 707-727.
- Rohwer, S. (1975). The social significance of avian winter plumage variability. *Evolution*, 29, 593-610.
- Scott-Phillips, T.C. (2006). Why talk? Speaking as selfish behaviour. In A. Cangelosi, A.D.M. Smith and K. Smith (Eds.), *The Evolution of Language* (pp. 299-306). Singapore: World Scientific.
- Seyfarth, R.M., Cheney, D.L., and Bergman, T.J. (2005). Primate social cognition and the origins of language. *Trends in Cognitive Science*, 9, 264-266.
- Shepher, J. (1971). Mate selection among second-generation kibbutz adolescents: Incest avoidance and negative imprinting. *Archives of Sexual Behaviour*, 1, 293-307.
- Shepher, J. (1983). *Incest: A Biosocial View*. New York, NY: Academic Press.
- Silk, J.B., Kaldor, E., and Boyd, R. (2000). Cheap talk when interests conflict. *Animal Behaviour*, 59, 423-432.
- Tallerman, M. (Ed.). (2005). *Language Origins: Perspectives on Evolution*. Oxford: Oxford University Press.
- Tinbergen, N. (1952). "Derived" activities; their causation, biological significance, origin, and emancipation during evolution. *Quarterly Review of Biology*, 27, 1-32.
- Tinbergen, N. (1963). On the aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20, 410-433.
- Ulbaek, I. (1998). The origin of language and cognition. In J.R. Hurford, M. Studdert-Kennedy and C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 30-43). Cambridge: Cambridge University Press.
- West, S.A., Griffin, A.S., and Gardner, A. (2007). Social semantics: Altruism, cooperation, mutualism and strong reciprocity. *Journal of Evolutionary Biology*, 20, 415-432.
- Whitfield, D.P. (1987). Plumage variability, status signaling and individual recognition in avian flocks. *Trends in Ecology and Evolution*, 2, 13-18.
- Wolf, A.P. (1970). Childhood association and sexual attraction: A further test of the Westermarck hypothesis. *American Anthropologist*, 72, 503-515.
- Wolf, A.P. (1995). *Sexual Attraction and Childhood Association: A Chinese Brief for Edward Westermarck*. Stanford: Stanford University Press.
- Wray, A. (Ed.). (2002). *The Transition to Language*. Oxford: Oxford University Press.
- Zahavi, A. (1975). Mate selection: A selection for a handicap. *Journal of Theoretical Biology*, 53, 205-214.
- Zahavi, A., and Zahavi, A. (1997). *The Handicap Principle: A Missing Piece of Darwin's Puzzle*. Oxford: Oxford University Press.