To appear in J. Wells (ed.), Social Information Transmission and Human Biology. London: Taylor & Francis.

Language: Costs and benefits of a specialised system for social information transmission

Daniel Nettle

Psychology, Brain and Behaviour

Henry Wellcome Building

University of Newcastle

Newcastle

NE2 4HH, UK

E-mail: daniel.nettle@ncl.ac.uk

Abstract

Language is often thought of as the crowning human adaptation, the one that allowed *Homo sapiens sapiens* to conquer the globe. The assumption underlying such ideas is that verbal transmission of information provides unalloyed benefits, by reducing the costs of learning about the environment. However, this raises the question of why no other species has discovered such a good trick. I argue that verbal transmission is only likely to be adaptive in a restricted range of circumstances. Even then, it cannot be exclusively relied on, and it causes problems of deceit and instances of maladaptation. We should expect natural selection to have made us discriminating evaluators of verbal information who ultimately trust the evidence of our senses. Nonetheless, once language has become widespread, it can increase human adaptability, by increasing the efficiency of individual learning.

1. The great leap forward?

In language, human beings have a unique system for the cheap, reliable exchange of learned information. This sets them apart from all other species. A number of other sophisticated information exchange systems have been documented in nature, but each is limited to a particular specialised domain, such as predator types for vervet monkeys (Seyfarth, Chenery, & Marler, 1980), food sources for bees (Von Frisch, 1967), or clan membership in cetaceans (Rendell & Whitehead, 2003). Human language, on the other hand, through the arbitrary and learned nature of its vocabulary, and the combinatorial properties of its syntax, allows an unlimited number of messages about an unbounded array of topics to be exchanged. Language is underlain by specialised neural structures, and thus appears to be a human-specific adaptation for generalised information exchange (Pinker & Bloom, 1990). I will call such exchange verbal transmission, and henceforth assume verbal transmission to be the primary function of language. Thus, henceforth, I will refer to language as a set of capacities, and verbal transmission as the activity that those capacities allow.

It is not difficult to identify putative benefits of verbal transmission. Indeed, accounts of the evolution of language have been more concerned with the origin of the necessary neural and morphological mechanisms than bothered by the question of what the adaptive value of the system might be – as if the latter were self-evident. Thus, for Pinker and Bloom (1990: 712):

> [There is] an obvious advantage to being able to acquire… information second-hand: by tapping into the vast reservoir of knowledge accumulated by other individuals, one can avoid having to duplicate the possibly time-consuming and dangerous trial-and-error process that won that knowledge.

Many accounts of human evolution have assumed or implied that the origin of language allowed the explosion in human numbers and dispersal that began some time before 50 thousand years ago, and led to the colonisation of every continent of the globe (Klein, 2000). It is certainly true that humans produce sophisticated and cumulative technological and social traditions, and that in virtue of these, they are able to successfully inhabit a vast range of environments, and gradually increase their exploitation of those environments. This is only possible because of their ability to learn culturally, in the broad sense of by transmission of information between individuals (Boyd & Richerson, 2004; Henrich & McElreath, 2003). Verbal transmission is a key component of cultural abilities, since it allows information to be encoded and transmitted with relatively high fidelity and low cost.

However, the advantages of verbal transmission immediately raise what is known as the 'why not baboons?' problem (Henrich & McElreath 2003: 126). If verbal transmission is so fantastically adaptive, why has only one of several highly social, relatively encephalised, group-living primates evolved it? Why, in short, is such a good trick so rare[1]? Answers to this question that rely exclusively on the contingent lack of availability of the relevant neural or articulatory machinery do not convince. For one thing, highly complex adaptations such as they eye have originated several times independently in nature, despite their reliance on rare and specialised tissues. For another, such a view would imply a Whig view of history, as if evolution were always trying to get to language, but somehow most species were in limbo waiting for the necessary mutations to arise.

In this paper, instead, I argue that reliance on verbal transmission has costs as well as benefits, costs active both in the ancestral past and today. This means that it can only invade a non-verbal population under restricted circumstances. These restricted circumstances must have, at some point, obtained for a prolonged period during human evolution, and during this period, the psychological adaptations for language were put in place. Once established, language clearly increased mean fitness and human adaptability, as the explosion of humans across the world attests. However, the question of exactly how adaptability was increased is not as simple as it may seem, as I shall show below.

In section 2, I review the costs and benefits of verbal transmitting information rather than representing it genetically or producing it by individual learning. I use the

[1] Cultural variation of a basic sort is relatively common in nature (Boyd & Richerson, 1996; Heyes & Galef, 1996), with numerous examples of socially transmitted local foraging practices. However, such variation has two features that differ from verbal transmission in humans. First, it is brought about by performance of the behaviour itself, not by conversion of information about the behaviour into another format (words) and back again. Second, the behavioural traditions are not cumulative, in the sense of leading to progressively more complex behaviours that no individual could have come up with alone.

considerations developed in section 2 to develop a formal model of the evolution of verbal transmission in section 3. In section 4, I consider how verbal transmission increases human adaptability, and section 5 presents the overall conclusions.

2. Costs and benefits of language

Benefit 1: Flexibility

Evolutionary theorists have addressed the question of when evolution should favour reliance on learning rather than provide behavioural strategies that are encoded genetically (Bergman & Feldman, 1995; Stephens, 1991). Fixed, innate strategies have the advantage of being available reliably whatever the contingencies of learning history happen to be, and are potentially available immediately at birth. Mutation notwithstanding, innate programmes produce the optimal behaviour more reliably and cheaply than learning, which involves more extra stages both inside and outside the organism. In general, these benefits will mean that behaviours will be encoded genetically unless there are counter-acting advantages. Defensive reflexes, crying and suckling are innate for these reasons.

Reliance on learned information can evolve where the rate of environmental change, and thus, change in the behaviour which is locally optimal, is too rapid for selection to assemble innate adaptations to produce the behaviour directly (Stephens, 1991). The disadvantages in terms of learning time and the risk of mis-learning are outweighed by the more rapid tracking of environmental variation in time or space. However, this leaves open the question of what kind of learning will be favoured. In general, there are two possibilities. In individual learning, the organism bases its behaviour on its own history of interaction with the environment. In social transmission, locally optimal strategies are passed from one individual to another either by imitation, or by encoding them in language[2]. Verbal transmission is thus a subset of social information transmission.

[2] There are also intermediate cases between individual and social learning, such as social enhancement of individual learning (see Boyd & Richerson, 1996; Heyes & Galef, 1996). This occurs where the activity of older animals increases the probability that naïve ones will manage to learn a particular behaviour, for example, by causing them to be in particular locations. Many animal cultural traditions belong to this weaker class than fully symbolic, cumulative human social information transmission. For

Benefit 2: Reduced learning costs and increased fidelity

Let us assume, then, that reliance on some kind of learning will be favoured when the environment changes too fast for selection to track the variation genetically. This raises the issue of what factors favour individual learning versus social, and particularly verbal, transmission. It has generally been assumed, as in the Pinker and Bloom quotation reproduced in section 1, that the chief advantage of verbal transmission is that it reduces the cost, in time or energy, of acquiring information. It is much quicker and safer to be told which of two plant species is toxic than to have to discover this by experiment. A benefit of verbal transmission will thus be cost reduction.

A second plausible benefit of verbal forms of information transmission is that fidelity of knowledge about the environment is increased. For example, an environment may contain a species of which 30% of examples are toxic. A simple statistical power calculation shows that the number of individual experiments required for there to be a 95% chance of finding out that this plant can be toxic is 9. This may be too many, in terms of risk of exposing oneself to a fatal level of toxin. Nine trials could also accumulated by observation, but encoding the information verbally avoids the unnecessary repetition of eating trials by new individuals who have not made enough observations to be sure. The information can be handed down through time and out across space at minimal extra cost.

Cost 1: Reliance on socially transmitted information

The discussion so far has focussed on the clear benefits to individuals of receiving verbal information. However, there are potential costs too. Reliance on such information, as Emile Durkheim first pointed out nearly one hundred years ago, binds human beings into an immense web of cooperation, stretching not only into space but through time as well (Durkheim, 1982: 248). Cooperation can bring adaptive benefits, but a huge literature in behavioural biology has concentrated on the difficulty of making cooperation evolutionarily viable (for some classic sources, recent developments and reviews, see Axelrod & Hamilton, 1981; Boyd & Richerson, in

---

clarity of argumentation in this chapter, I contrast pure verbal transmission with pure individual learning, though recognising that there are in fact intermediate possibilities.

press; Doebili, Hauert, & Killingback, 2004; Dugatkin, 1997; Hamilton, 1964; Panchanathan & Boyd, 2003; Trivers, 1971). The essential problem is that cooperation involves taking a personal cost (or at the very least, an opportunity cost) in order to reap a larger benefit. This leads to the evolution of strategies that take the benefit of cooperation without paying the costs. Such free-riding strategies soon invade a cooperating population, leading to the extinction of cooperation, since however badly a free-rider does, a co-operator in the same population will always do even worse.

In the evolution of signalling, this problem is particularly acute, because signals can be faked and thus are potentially misleading. Individuals maximise their own fitness, not that of the population, and where individuals have conflicting interests, it is often adaptive for them to withhold information or deceive their rivals. As long as signals are cheap to produce, as language is, then under many conditions, signals will become worthless as the system is invaded by misleading information (Dawkins & Krebs, 1978; Maynard-Smith, 1982; Silk, Kaldor, & Boyd, 2000).

Humans have somehow overcome this problem, at least often enough for language as a system to be viable. How they might have done so is discussed below. Certainly, humans show a general willingness to provide benefits to each other that is greater than that which seems optimal from an immediate fitness perspective (Boyd & Richerson, in press; Gintis, Bowles, Boyd, & Fehr, 2003). In general, to rely on information cheaply signalled by conspecifics is to expose oneself to the possibility that that information is false. To provide truthful information to conspecifics where there is any cost at all of doing so, is a cooperative act. The development of language would thus be unlikely to be adaptive in a context where very extensive cooperation had not become prevalent, and such cooperation is rare in nature. Moreover, although the problem in general may have been overcome in humans, experience shows that there is still residual deception of individuals by others. Thus, the cost of relying on verbal information is still operative, since it might have been provided maliciously, or transmitted sincerely by someone who got it from someone who provided it maliciously, etc. Thus, in the model in section 3, it is necessary to include a parameter that represents the probability of information being reliable.

There are several conditions which allow the problem of honest signalling to be overcome. Where individuals are closely related, they share sufficient genetic interests for a higher level of cooperation to be maintained than where they are unrelated (Hamilton, 1964). This principle accounts for the existence of cheap but honest signalling about the whereabouts of resources in social insects, where levels of relatedness are high. Individuals that encounter each other frequently can develop cooperation based on reciprocity, with one providing a benefit to the other in return for the reverse at a later point in turn (Trivers, 1971). Honest signalling can arise in this manner (Silk et al., 2000). However, such reciprocal altruism is easily disrupted by free-riders, and unstable in sizable groups (Boyd & Richerson, 1988). For these reasons it is thought to be quite rare in nature, and insufficient to account for the general human propensity to cooperate.

In sizeable groups, it is more plausible that cooperation is maintained by indirect reciprocity than by dyadic reciprocation. Indirect reciprocity refers to situations where non-cooperators become publicly known as such, and are excluded from future cooperation not just by the individuals they failed to cooperate with, but by everyone in the group (Panchanathan & Boyd, 2003). The maintenance of indirect reciprocity is easiest when there is a mechanism of social information transmission about non-cooperators, and it has been argued that one of the primary functions of language is to exchange social information (gossip) of this kind (Dunbar, 1996). However, this obviously poses a second-order problem; how to know that gossip is reliable and not malicious (Power, 1998). Cooperation is a *prerequisite* for language to be reliable, so if language is a prerequisite for indirect reciprocity, and indirect reciprocity sustains human cooperation, we have a paradox. In small communities, indirect reciprocity could be sustained by observational learning, and then language and more extensive cooperation bootstrap off each other. Though indirect reciprocity is theoretically plausible, it cannot be assumed that it is universally stable, and thus the stability of cooperation cannot be taken for granted.

Two other factors have been proposed for the maintenance of cooperation. One is punishment; if individuals punish non-conformists then cooperation, or indeed any other behaviour, can be stabilised within sizeable groups (Boyd & Richerson, 1992). This too raises a second-order problem. If individuals have to punish at a cost to themselves, then there will arise free-riders who benefit from being in a group with punishment, but do not contribute to the punishment activity. Soon punishment goes extinct as individuals who skimp on the cost of punishing out-compete punishers. Panchanathan and Boyd (2003) have shown that the second-order problem can be solved where punishment takes the form of the withholding of benefits, rather than the inflicting of costs, since in this case the costs of punishment are negative and there is no temptation to skimp on them. In this scenario, punishment and indirect reciprocity are two sides of the same coin. However, even in this scenario, though cooperation is an evolutionarily stable strategy, so is a population of non-cooperators. Cooperators can only invade a population composed of non-cooperators if they are given an initial advantage by being allowed to preferentially assort with each other or their kin when they are rare. This is a common feature of many models of the evolution of cooperation (see e.g. Nettle & Dunbar, 1997).

There is a final factor which has been argued to potentially account for the propensity of human beings to cooperate so widely – group selection. Group selection occurs when the total population is divided into interacting groups which persist or go extinct at different rates according to the properties of the individuals within them. Group selection can in principle favour the proliferation of traits that help the whole group to survive even at cost to the individual. However, the conditions for group selection to be important are extremely restrictive. For one thing, selection between individuals within each group will tend to reduce the proportion of group altruists, and since the extinction rate of individuals is much faster than that of groups, this will be the dominant force. Moreover, migration amongst groups reduces group-level variation, which is required for selection to operate. Thus group selection is generally swamped by individual-level selection, and the result is the persistence of no more cooperation than would be predicted by assuming groups did not exist.

However, the argument has been made that group selection based on groups sharing cultural characteristics is plausible (Soltis, Boyd, & Richerson, 1995). Because human societies enforce conformity in cultural practices, by punishment or conformist adoption of norms, group traits can be maintained homogenously even in the face of a flow of migrants, or differential reproductive success of individuals within the group. As long as cultural groups sometimes die out or disperse, which is true for example of

New Guinea clans (Soltis et al., 1995), and as long as the probability of this is related to the proportion of pro-social behaviours within the group, group selection will be operative. However, it is still likely to be a weak force compared to individual-level selection, and no-one claims that humans are unconditional pro-group altruists. Furthermore, the possibility of cultural group selection depends on the prior existence of cultural, perhaps including linguistic, transmission of norms.

The problem of human cooperation is far from solved, and it is beyond the scope of this paper to discuss it further. However, the brief review given so far allows a number of conclusions relevant to the current thesis that language is not unconditionally adaptive. These conclusions are:

1. The cooperativeness of unrelated individuals with one another should not be assumed. For cooperation to persist in sizeable groups, one or several of a number of specific and rare conditions must obtain. Widespread cooperation is rare in nature.
2. The benefits of verbal transmission depend on the stability of widespread cooperation, otherwise the information received verbally, which is cheap, fakeable and often not immediately verifiable, would not be reliable.
3. Several of the factors which might explain human cooperativeness, such as indirect reciprocity and cultural group selection, are dependent on or at least facilitated by the existence of language, thus causing a chicken-and-egg problem. It may be that language began to be used in small groups characterised by high levels of genetic relatedness and limited opportunity to defect, but that once in place, language and culture potentiated the expansion of groups, which further enhanced the selective benefits of using verbally transmitted information, and so on in a feed-forward loop.
4. Even in fully modern populations, humans are not unconditional cooperators, and thus a potential drawback of using verbal information is that it may be unreliable. This should be considered in any model of the evolution of language.

Cost 2: Environmental heterogeneity and change

Under benefit 1, I argued that an advantage of using learned over innate information was that it allowed the tracking of environmental change too rapid for selection to follow. This is certainly an advantage of language relative to innate information. On the other hand, a cost of using linguistically transmitted learned information rather than individually learned information is that it allows slower response to environmental change. To see why this is the case, consider information about a foraging strategy that is successful in the local environment. If this information is acquired by the individual in the place where he currently is, there will be a costly learning process, but on the other hand, the information will be guaranteed to be locally appropriate and up to date. If the information is acquired verbally, then the source of the information will have either learned it individually himself in a previous time step, or in turn acquired it verbally from someone else. If he acquired it verbally, then the person from whom he acquired it will in turn have either learned it individually or acquired it verbally, and so on ad infinitum.

The longer the chain of verbal links is before someone is reached who actually learned the information first hand, then the greater probability is that the environment in which the information was originally learned has changed or is different from the one at the current location. If environmental change is very fast, or equivalently, if the habitat is spatially very heterogeneous, then individual learning is favoured over social transmission. This is because, despite the costs of individual learning, it is better to learn in the current place or time than be given hand-me-down information that is outdated or locally non-optimal. This result can be shown in a wide variety of models (see section 3), and leads to the general conclusion that social transmission is only adaptive where the environment changes too fast to be tracked by selection on genes, but not so fast that constant individual learning is required (Boyd & Richerson, 1985; Henrich & McElreath, 2003). The late Pleistocene was characterised by dramatic climate fluctuations over the order of a few decades – fast in evolutionary terms, but long relative to human lifespan. This leads to the influential idea that human reliance on language and culture is an adaptation to Pleistocene climatic variability (Boyd & Richerson, 2004).

Cost 3: The Chinese whispers problem

As described above, the more verbal transmission there is in the population, the longer on average the chain of verbal information links is before someone is reached who actually acquired the information first hand. This leads to a puzzling result (Boyd & Richerson, 1995; Rogers, 1988). As the proportion of reliance on verbal information (versus individual learning) increases, the value of that information drops, and the fitness of verbal learners decreases. This is shown analytically in section 3, but to see why it must be the case, consider the extremes. When verbal learning is extremely rare, then any information acquired verbally will almost certainly have been learned individually by the next person in the chain, and thus is likely to be up to date and locally apt. On the other hand, consider the case where verbal learning has gone to fixation in a population. Now, everyone simply relies on verbal accounts of what has been done before, and no new learning brings better information into the system. This becomes like the old parlour game of Chinese Whispers; a sentence is passed round and round a circle of people, gradually becoming more distorted. In this case, it becomes distorted not because of the fidelity of language, but because as the environment changes, the information becomes less and less useful.

Two things follow from the Chinese Whispers problem. One is that 100% verbal transmission of information is not adaptive (Rogers, 1988). There will instead always be a mixed equilibrium of some verbal transmission and some individual learning (see section 3). Someone somewhere has to actually be testing things out, otherwise unreliable information is just recycled. Second, the mean fitness of a population at equilibrium can never be increased by the adoption of verbal transmission alone. This is because reliance on verbal information increases to the point where the payoff of the verbal strategy is identical to that of individual learning (see section 3). This is a very important point, since those who argue that it was the adoption of language that allowed humans to dramatically expand their numbers across the globe are assuming that using language enhanced mean fitness. It may well be that it did, but for that to be the case, it must have done more than just eliminate learning costs. I consider the question of what else it does for humans in section 4, below.

Summary of costs and benefits

To summarise section 2, verbal transmission has a number of benefits and costs. It is more flexible than reliance on innate information, though not as flexible as individual learning. We should thus expect it to be favoured in situations where the environment changes rapidly, but not too rapidly. It saves the costs of individual learning. However, on the cost side, it makes organisms reliant on information that is fakeable and potentially incorrect. Thus it is only likely to be adaptive either where individuals are closely related or where widespread cooperation has somehow become established. Even then, there is a residual risk of deception. Moreover, because of the Chinese whispers problem, the whole population can never be exclusively reliant on verbally transmitted information. There must always be a component of individual learning updating information within the system. Finally, by reducing learning costs alone, verbal transmission can never increase the mean fitness of a population. Any absolute increment in fitness and adaptability due to language must come from another source.

The inventory of costs and benefits laid out in this section leads naturally to the statement of a more formal model of the evolution of verbal transmission. This model, which draws heavily on those of Rogers (1988) and Boyd and Richerson (various papers, especially 1995), allows more precise predictions to be made about the general conditions for the evolution of language.

3. Evolution of verbal transmission: A formal model

Let us assume that there is a temporally variable environment with an infinite number of states. The organisms must forage in that environment as best they can. There is a skill available that increases foraging efficiency, so that unskilled individuals have fitness w, but skilled individuals have fitness $w + b$. Each generation there is a probability u that the environment switches to another state in which the old skill is no longer useful, though there exists a new skill for the new environment that gives fitness $w + b$.

We will consider two genotypes with different learning rules. *Individual learners* acquire the skill individually, with probability d and cost c. The parameter d is the efficiency of learning; where it is low even prolonged trial and error does not guarantee that the optimal skill will be found. The parameter c is the cost, in time or energy, of learning something for oneself. *Verbal learners* obtain their information verbally from another in the community, chosen at random. There is no learning cost for them (we can make the model more complex by making verbal learning have a cost, but since we generally assume that the costs of verbal learning are less than those of individual learning, this case can be ignored for now). More complex genotypes that use a combination of individual and verbal learning would be more realistic, but the use of the two simple genotypes allow the model to be simple, and does not affect the general conclusions. Similarly, a more complex model would also allow possibilities intermediate between individual and verbal learning, such as social enhancement of individual learning, but for simplicity, the only comparison here is between the two extreme alternatives.

Consider a population with a frequency p of verbal learners and $(1-p)$ of individual learners. The fitness of the individual learners is:

$$w_I = w + db - c \quad \text{(E1)}$$

Verbal learners acquire behaviours that were originally individually learned at some earlier time (through a chain of intermediaries). The probability that the behaviour was originally learned one generation ago would be $(1-p)$. The probability that it was learned a generation before that would be $p(1-p)$, and the generation before that, $p^2(1-p)$. Thus in general, the probability that the behaviour was originally learned r generations ago[3] is $p^{(r-1)}(1-p)$.

The fitness of a verbal learner will be $w + b$ if the skill they learn about verbally is the optimal one for the current environment, and w if it is not. There is some chance, e that they will be fed misleading or malicious information. The probability of the information being sincere is thus $(1-e)$. The information will also only be optimal if there has been no environmental perturbation since the behaviour was originally learned. The probability that there has been no perturbation in r generations is $(1-u)^r$. Therefore the fitness of a verbal learner will be given by E2.

$$w_v = w + (1-e)b\sum_{r=1}^{\infty} p^{r-1}(1-p)(1-u)^r \quad \text{(E2)}$$

E2 can be rearranged to give E3.

---

[3] It is assumed here that natural selection is weak, that is, cultural change is much faster than genetic change, so that its effect on p can be ignored in this calculation.

$$w_v = w + b(1-e)(1-p)(1-u)\sum_{r=1}^{\infty} p^{r-1}(1-u)^{r-1} \qquad \text{(E3)}$$

Using the formula for the sum of an infinite geometric series, E3 gives E4.

$$w_v = w + b \bullet \frac{(1-e)(1-p)(1-u)}{1-p(1-u)} \qquad \text{(E4)}$$

First, let us derive the Chinese whispers problem, by a simple combination of parameters where verbal information is perfectly honest, the environment changes moderately fast, and individual learning is moderately costly (see Figure 1 and legend). The result shown in figure 1 is that, given we have assumed that verbal information is generally reliable, and environmental change is quite slow, verbal learners have a large fitness advantage when rare. However, this advantage diminishes as their numbers increase. At fixation, their fitness is less than that of individual learners, and as a consequence, the evolutionary equilibrium is always a mixture of strategies[4]. There is another important consequence; at equilibrium, the mean fitness of an individual in the population, regardless of which strategy they follow, is exactly the same as if everyone was an individual learner. Verbal learning has become widespread, but not increased the average fitness in the population.

Now consider the conditions for a rare verbal learner to become numerous in a population of individual learners. Thus we set *p* close to zero. Verbal learners will increase in numbers where $w_v > w_I$, that is where E5 is satisfied.

$$w + b(1-u)(1-e) > w + db - c \qquad \text{(E5)}$$

E5 can be arranged to give E6.

$$b(1-u)(1-e) > db - c \qquad \text{(E6)}$$

[4] In this version of the model, a mixture of genotypes, but a single genotype that relied on a mixture of verbal and individual learning would also be an ESS.

E6 defines an area of the parameter space where verbal transmission can become more common when rare. Some illustrative results are given in figure 2. The model shows that there are combinations of parameters where verbal transmission can spread, and many others where it cannot. In general, other things remaining constant:

1. As environmental change gets faster, or malicious information becomes more common, it gets more difficult for verbal transmission to spread.
2. As the efficiency of individual learning diminishes, verbal transmission becomes more likely to spread.
3. As the costs of individual learning increase, the fitness advantage of verbal transmission increases.

The various parameters are traded off against each other, so that, for example, where the cost of individual learning is low, then it is hard for verbal transmission to evolve even if verbal information is very reliable, but as the cost of individual learning increases, the reliability threshold required for verbal transmission gets lower (figure 2a). Similarly, if environmental change is very fast, then verbal transmission cannot spread even if the efficiency of individual learning is very low (figure 2b).

4. Language and complex cultural adaptations

We have seen that linguistic transmission can proliferate under a certain set of conditions, and equilibrate at the point where the mean fitness of individuals is exactly the same whether they use language or individual learning. However, the argument usually made is that language *increased* average human fitness, allowing the population to explode and prosper in new habitats all over the world. Boyd and Richerson (1995) have argued, correctly, that this must mean that language does more than just reduce the costs of learning that individuals could achieve on their own. For language or culture to increase human adaptability, they must actually allow better adaptations to be acquired than individual learning could ever achieve.

Language surely does this, not by replacing individual learning, but by canalising it. Language provides an enormous compression of information, and because it is categorical, a huge reduction of ambiguity. Verbal information is available as background for individual experiment. In skilled domains such as foraging, medicine or construction of technology, verbal information interacts dynamically with individual learning, pre-guiding it to fruitful regions, and pre-warning it of possible pitfalls. This allows more reliable discovery of complex local solutions in a multi-dimensional fitness landscape where individual trial and error might get stuck in a valley. Individuals add their learning to the existing structure and pass it back into the mix, so that cognition becomes distributed across many individuals (Hutchins, 1995). Just as adding neurons to a neural network increases its computational power, so adding individuals to a community increases its cultural learning potential. Note again, though, that this relies on verbal transmission never being a pure strategy, but individuals retaining discriminating ultimate reliance on their individual experience[5].

Language may also have benefits internal to the individual. There has been much interest within linguistics in the notion that language is a system of internal representation as much as inter-individual communication (Hauser, Chomsky, & Fitch, 2002). Because linguistic representations are discrete and systematically combinable, they can be used internally for processes of logical reasoning and off-line planning of behaviour. This does not mean that the language of mental representations is narrowly tied to any particular natural language, but rather that the evolution of a discrete combinatorial communication system made additional representational possibilities available to internal as well as communicative processes, which in turn increased the efficiency and scope of individual learning.

Language as a system of representation, and distributed cognition, mean that cultures can acquire adaptations that no non-linguistic individual could ever have come up with. Put together they allow cumulative and increasingly sophisticated knowledge-based local adaptations. They unite powerfully when cultures develop external means of information storage, in tallies, painting, writing and ultimately computers. These

[5] Indeed, if conformism is very strong within a group, then adding more individuals is not guaranteed to lead to the discovery of better adaptations. The viability of cultural group selection demands high conformism, but the efficiency of distributed cognition demands scepticism and independent-mindedness. These conflicting influences can be seen in many cultural traditions, including of course science itself.

developments increase the potential for discovery of good local adaptations by a further order of magnitude (Donald, 1991).

4. Conclusion: Why language is rare.

We have seen that there are great rewards in store for any species that gets an efficient system of social information transmission running widely. As so often argued, this may be the key to the human demographic and cultural explosion. Language would allow this not by replacing individual learning but by allowing it to become more efficient and better guided. But if the fitness benefits are so overwhelming, why is language so rare? In short, why not baboons?

We have seen that the rarity of language is not the simple contingency of the right mental machinery not having been widely available. Selection is a powerful mechanism for producing adaptations, and where there is a fitness payoff, it has had no trouble producing the discrete alarm calls of vervet monkeys or the waggle dance of bees, to cite but two examples. Instead, we have seen that verbal transmission would only be adaptive in a restricted set of circumstances; where environmental change is fast enough but not too fast; where individual learning is not too cheap or efficient; and above all, where cooperation is widespread. It is probably rare for these circumstances to prevail, but once they do, the existence of language reinforces its benefits, for example by allowing cooperation to be policed and group norms to be agreed upon (Boehm, 1996). Thus language is a good illustration of the path-dependent nature of evolution.

Moreover, the initial transition to widespread language would not immediately lead to the human explosion, since, where language simply reduces learning costs, we have at seen that at equilibrium it does not increase average fitness. Rather, the transition to language to save learning costs made possible later cognitive adaptations whose benefit was to make individual learning more sophisticated. Distributed cognition leads to cumulative increases in the sophistication of learned adaptations, which gives human cultures their initially slow but undeniably directional increase in overall complexity.

However, even with all these benefits in place, language is a costly business. Partly because of language, people can be persuaded to behave in ways detrimental to their fitness, and can end up with beliefs and values far from optimal in their environments. Even where cultural transmission is generally adaptive, it can lead to many instances of groups behaving in a collectively maladaptive way, especially if conformism is strong (Boyd & Richerson, 1985). In America in the first half of the twentieth century, there was a craze for radioactive drinks, cures and spas. People spent considerable amounts of money on irradiating themselves, at least in part because of verbal transmission. A Dr. C.G. Davis had written in the *American Journal of Clinical Medicine* that 'Radioactivity prevents insanity, rouses noble emotions, retards old age, and creates a splendid youthful joyous life'.

This maladaptative cultural practice, which did not die out until the late 1950s, could never have become prevalent by trial and error learning, since it had no positive effects whatever. Deceit and credulity must have been recurring problems in human evolution, and continue to be so today. As a result, we have sophisticated intuitions about social cheaters and deception, and are not cultural dupes (as some social constructionists seem to imply), but generally sceptical evaluators of verbal information who ultimately trust most in the evidence of our senses. Language is the crowning human adaptation, but like any other adaptation, it had costs as well as benefits, and is only adaptive in relation to a particular set of circumstances.

Figure 1. The Chinese whispers problem. Using the model outlined in the text, with, for simplicity, e at 0, w and b at 1, u at 0.1, d at 0.9, and c at 0.2. p is allowed to vary from 0 (no verbal learners) to 1 (all verbal learners). Whereas the fitness of individual learners is constant, fitness of verbal learners declines as their frequency increases, such that at fixation their fitness would be at background level w. As a result, the evolutionary equilibrium is always a mixture of strategies with proportions determined by the point of crossing of the two lines (in this instance, about 75% verbal learning). Making verbal information less reliable or the speed of environmental change faster shifts the equilibrium point to the left. Making individual learning more costly or less reliable shifts it to the right.

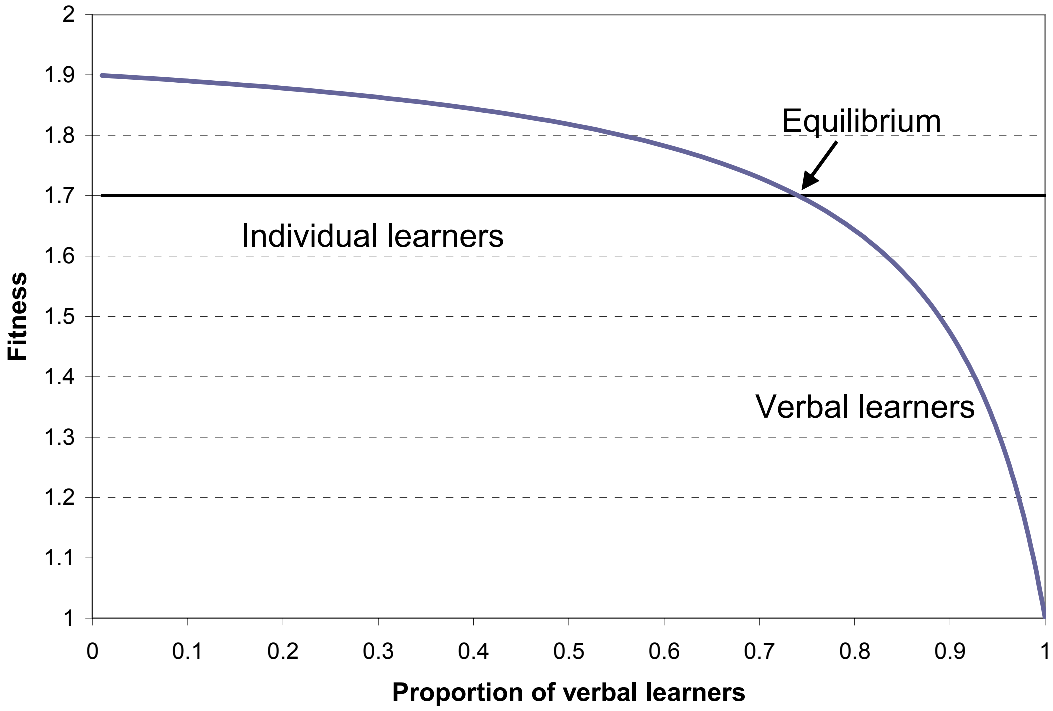
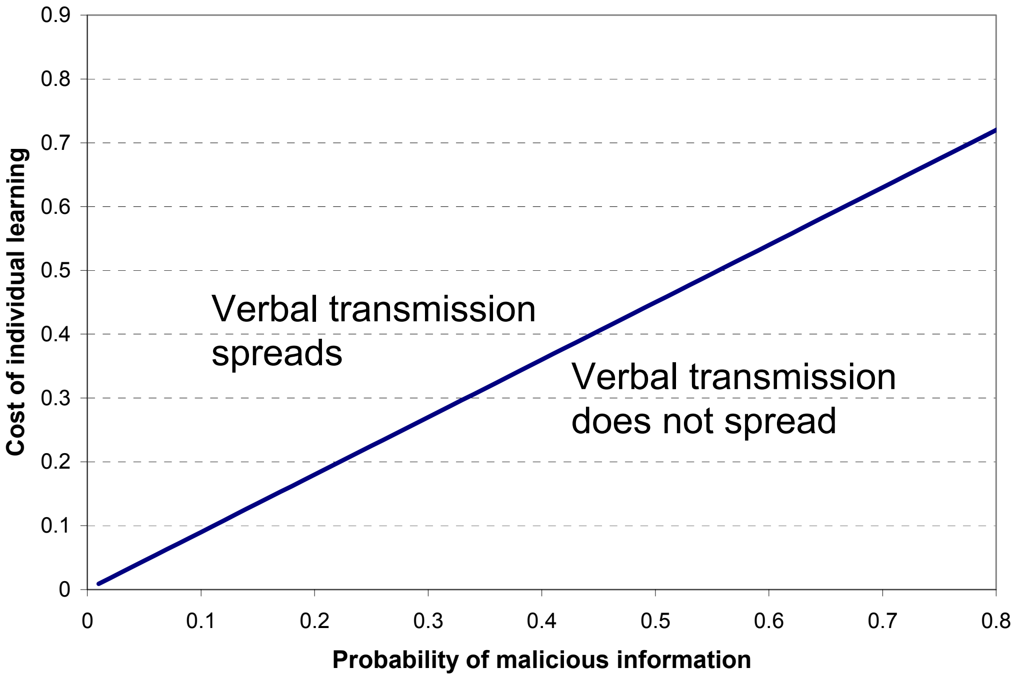


Figure 2. Regions of parameter space where verbal transmission can and cannot spread when it is rare. *w* and *b* are set at 1.

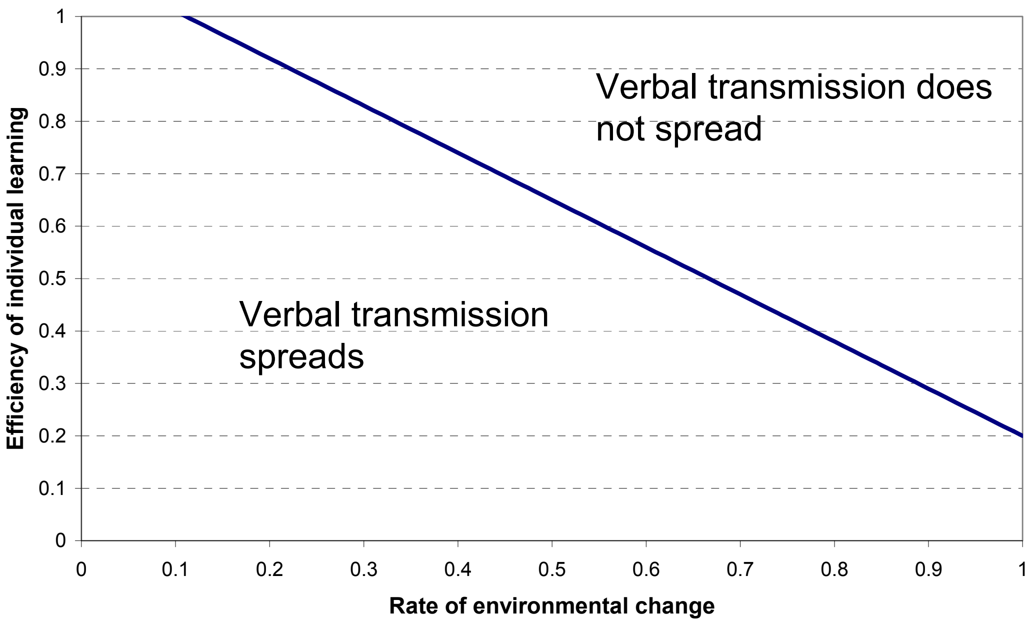
(a) *u*=0.1 and *d*=0.9. Verbal transmission can spread if the cost of individual learning is high, and/or the probability of malicious information is low.

(b) *e*=0.1 and *c*=0.2. Verbal transmission cannot spread if environmental change is very fast and/or individual learning very efficient.

(a)



(b)



References

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*, 1390-1396.

Bergman, A., & Feldman, M. W. (1995). On the evolution of learning: Representation of a stochastic environment. *Theoretical Population Biology, 48*, 251-276.

Boehm, C. (1996). Emergency Decisions, Cultural Selection Mechanics, and Group Selection. *Current Anthropology, 37*, 763-793.

Boyd, R., & Richerson, P. J. (1985). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *Journal of Theoretical Biology, 132*, 337-356.

Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology, 13*, 171-195.

Boyd, R., & Richerson, P. J. (1995). Why does culture increase human adaptability? *Ethology and Sociobiology, 16*, 125-143.

Boyd, R., & Richerson, P. J. (1996). Why culture is common but cultural evolution is rare. *Proceedings of the British Academy, 88*, 73-93.

Boyd, R., & Richerson, P. J. (2004). *Not by Genes Alone*. Chicago: University of Chicago Press.

Boyd, R., & Richerson, P. J. (in press). Solving the puzzle of human cooperation. In S. Levinson (Ed.), *Evolution and Culture*. Cambridge, MA: MIT Press.

Dawkins, R., & Krebs, J. R. (1978). Animal signals: Information or manipulation? In J. R. Krebs & N. B. Davies (Eds.), *Behavioural Ecology: An Evolutionary Approach* (1st ed., pp. 282-309). Oxford: Blackwell.

Doebili, M., Hauert, C., & Killingback, T. (2004). The evolutionary origin of cooperators and defectors. *Science, 306*, 859-862.

Donald, M. (1991). *The Origins of the Modern Mind*. Cambridge, MA: Harvard University Press.

Dugatkin, L. A. (1997). *Cooperation Among Animals: An Evolutionary Perspective*. New York: Oxford University Press.

Dunbar, R. I. M. (1996). *Grooming, Gossip and the Evolution of Language*. London: Faber.

Durkheim, E. (1982). *The Rules of the Sociological Method*. London: Macmillan.

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior, 24*, 153-172.

Hamilton, W. D. (1964). The genetical evolution of social behavior. Parts I and II. *Journal of Theoretical Biology, 7*, 1-52.

Hauser, M., Chomsky, N. A., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*, 1569-1579.

Henrich, J., & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology, 12*(3), 123-135.

Heyes, C., & Galef, B. G. (1996). *Social Learning in Animals: The Roots of Culture*. San Diego: Academic Press.

Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Klein, R. G. (2000). Archeology and the evolution of human behavior. *Evolutionary Anthropology, 9*, 17-36.

Maynard-Smith, J. (1982). Do animals convey information about their intentions? *Journal of Theoretical Biology, 97*, 1-5.

Nettle, D., & Dunbar, R. I. M. (1997). Social markers and the evolution of reciprocal exchange. *Current Anthropology, 38*, 93-99.

Panchanathan, K., & Boyd, R. (2003). A tale of two defectors: The importance of standing for the evolution of indirect reciprocity. *224*, 115-126.
Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences, 13*, 707-726.
Power, C. (1998). Old wives' tales: The gossip hypothesis and the reliability of cheap signals. In J. R. Hurford & M. Studdert-Kennedy & C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 111-129). Cambridge: Cambridge University Press.
Rendell, L. E., & Whitehead, H. (2003). Vocal clans in sperm whales (Physeter macrocephalus). *Proceedings of the Royal Society of London Series B-Biological Sciences, 270*, 225-231.
Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist, 90*, 819-831.
Seyfarth, R. M., Chenery, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour, 28*, 1070-1094.
Silk, J. B., Kaldor, E., & Boyd, R. (2000). Cheap talk when interests conflict. *Animal Behaviour, 59*, 423-432.
Soltis, J., Boyd, R., & Richerson, P. J. (1995). Can group-functional behaviors evolve by cultural-group selection - an empirical test. *Current Anthropology, 36*, 473-494.
Stephens, D. W. (1991). Change, regularity, and value in the evolution of animal learning. *Behavioral Ecology, 2*, 77-89.
Trivers, R. L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology, 46*, 35-57.
Von Frisch, K. (1967). *The Dance Language and Orientation of Bees*. Cambridge, MA: Harvard University Press.