Review TRENDS in Ecology and Evolution Vol.xxx No.x
Full text provided by www.sciencedirect.com

Selective scenarios for the emergence of natural language

Szabolcs Számadó[1] and Eörs Szathmáry[1,2,3]

[1] Collegium Budapest (Institute for Advanced Study), Szentháromság u. 2, H-1014, Budapest, Hungary
[2] Biological Institute, Eötvös University, 1/C Pazmany Peter, H-1117, Budapest, Hungary
[3] The Parmenides Foundation, 16 Capo Perla, 57031, Capoliveri, Italy

The recent blossoming of evolutionary linguistics has resulted in a variety of theories that attempt to provide a selective scenario for the evolution of early language. However, their overabundance makes many researchers sceptical of such theorising. Here, we suggest that a more rigorous approach is needed towards their construction although, despite justified scepticism, there is no agreement as to the criteria that should be used to determine the validity of the various competing theories. We attempt to fill this gap by providing criteria upon which the various historical narratives can be judged. Although individually none of these criteria are highly constraining, taken together they could provide a useful evolutionary framework for thinking about the evolution of human language.

Introduction

The issue of the origin of human language is regarded by some as the hardest problem in science [1], partly because three timescales [those of phylogeny, ontogeny and learning (cultural transmission)] are involved. The problem is exacerbated by that fact that this transition is unique, given that it has occurred in only one lineage [2]. Human language is also a novel inheritance system [3], which has opened up the possibility for cumulative cultural evolution [4]; it has enabled the emergence of a complex society that rests on the negotiated division of labour and on the collaboration of large non-kin groups [5,6]. The issue of the origin of human language has thus provided fertile ground for speculation, and various alternative theories have been proposed (Box 1).

The main reason why this issue remains is that we know little not only about the first steps of the evolution of human language, but also the relevant anatomy and genetics of humans that first started to use it. To give a detailed list of the still-unresolved questions about language evolution is outside the scope of this review; however, the crucial ones include: in what context did language evolve? What was such communication about? What were the first words? What cognitive skills are sufficient and necessary for language to evolve? What are the neurobiological features of the brain that enable humans to acquire and process language so efficiently? What is the exact genetic background of human language? What part of human language is genetically determined (if any) and what parts are transmitted culturally? Finally, why is human language unique, that is, why did our ancestors use a conventional communication system instead of the traditional self-reporting signals used by most animals, and why do no other living species have a comparable means of communication?

These issues can be grouped according to whether they concern the relevant anatomy of early hominins or the relevant social structure, hence the context versus content of communication. Such a grouping relates to whether the evolution of human language was variation- or selection-limited, although a transition (such as the emergence of the genetic code, or eukaryotic sex) can be limited simultaneously by unlikely variation (e.g. evolution of novel cell-cycle proteins) and special selective requirements (e.g. favoring regularly alternating ploidy levels) [7] (see [2] for a discussion on major transitions).

Variation versus selection limitation

A transition is variation limited when the available genetic variation in the given lineage does not offer even a partial solution to the problem at hand, and it takes a considerable time (in evolutionary terms) for the necessary variation to arise. By contrast, a transition is selection limited if the necessary genetic prerequisites of a possible transition are present, but the given transition is not selected for as this would require a specific ecological or social context. For example, the origin of the eukaryotic cell is thought to have been variation limited [8].

Was the evolution of human language variation or selection limited? First, primates can learn rather well as evidenced by the cognitive capacities of some great apes [9], whereby they are able to learn a vocabulary of lexicon (be it a sign language or lexigrams) of up to several hundreds of words, and can use this in a flexible way, combining words in a novel way in novel situations; they also show some understanding of novel sentences [9]. These are impressive skills, even if their production ability is more limited compared with their understanding. Second, given that the chimp and human genomes (setting neutral variations aside) are so similar [10], it might be that relatively few (probably regulatory [11]) genetic changes were involved in the evolution of language (Box 2). Investigating the history of evolution shows that a highly complex organ (e.g. the eye) or behaviour (e.g. flight, sociality or reproductive division of labour) can evolve several times, provided that there is robust selection pressure for it; such examples also offer striking cases of evolutionary convergence [12]. This

Corresponding author: Számadó, S. (szamszab@ludens.elte.hu).

www.sciencedirect.com 0169-5347/$ – see front matter © 2006 Elsevier Ltd. All rights reserved. doi:10.1016/j.tree.2006.06.021

Glossary

Adaptationist approach: seeks to explain behavioural or morphological features of organisms in terms of adaptations to the living or non-living environment; that is, these features can be observed because they confer a fitness advantage to the organism. This assumption is the heart and soul of darwinian evolutionary theory, but it does not necessarily apply to all features.
Concordance: twins are concordant for a trait if either both or neither express the trait. Twins are discordant for a trait if it is exhibited by only one twin.
Conflict of interest: players rank differently the possible outcomes of an interaction.
Conventional communication system: a communication system using conventional signals (i.e. symbols).
Domain general cognitive skills: a cognitive skill that is not a specific adaptation to a specific task (such as linguistic performance, decision making, or resolution of social conflict); but is used in several or all different domains (such as memory).
Domain specific cognitive skills: cognitive skills that evolved in response to a specific problem, such as the homing instinct of pigeons or cheater detection in some primates (including humans).
Evolutionarily stable strategy (ESS): a strategy that, if all the members of the population play it, then no other mutant strategy can invade.
Equilibrium cost: cost of signals paid when all members of the population play the ESS.
Gestural communication: non-verbal communication that should not be confused with sign languages. Gestures confer additional information about the motivation, emotional state, feelings of the signaller, and can act as 'meta-signals' by altering the meaning of verbal communication (e.g. smiling, grinning or head-shaking).
Groundedness: the ability of the agents to relate the sign to the (external) referent by non-verbal means (e.g. by pointing).
Honest signalling: when signals reveal the relevant quality of the signaller to the receiver.
FLB: the faculty of language in the broad sense: that is, aspects shared by other faculties or that can be found in other organisms (e.g. voluntary control over signal production or vocal imitation and invention).
FLN: the faculty of language in a narrow sense: that is, unique aspects of language (recursion).
Major transitions: radical changes in evolution in the way that heritable information is stored, used and transmitted. Usually accompanied by the formation of higher level units of evolution. Local interactions and division of labour and/or a combination of functions also have an important role.
Naming game: a game in which agents have to evolve a shared set of conventions to name frequently observed features in their environment.
Non-adaptationist approach: an approach that argues that some behavioural or morphological features cannot be explained as being adaptations; instead some (perhaps many) features arise as byproducts of physical or biological constraints, or are neutral in terms of fitness.
Other-reporting signals: signals that carry information about the (living or non-living) environment; only a small fraction of animal signals fall into this category, most notably the 'dance' of honeybees and the alarm calls of vervet monkeys.
Selection-limited transition: the necessary pre-requisites of a possible transition are present, yet the given transition is not selected for, because to do so requires a special ecological or social context.
Self-reporting signals: signals that carry information about the state, motivation or future behaviour of the signaller; most signals used by animals fall into this category.
Sign languages: symbolic communication system built on visual instead of vocal signs. They are the visual equivalents of spoken languages, and should not be confused with gestures.
Spandrels: an unselected trait that is a byproduct of some other trait that is selected; a favourite example of non-adaptationists.
Variation-limited transition: a situation in which the available genetic variation in the given lineage does not offer even a partial solution to the problem at hand, and it takes a considerable time (in evolutionary terms) for the necessary variation to arise.

Box 1. Theories of language and language evolution

Theories of language and language evolution can be divided into two sets of hypotheses: a nativist versus empiricist account and a non-adaptationist versus adaptationist account, respectively (K. Smith, PhD Thesis, University of Edinburgh, 2003).

Language
The nativist paradigm argues that language capacity is a collection of domain-specific cognitive skills that is unique to humans and that is somehow encoded into our genome. Perhaps the most famous proponent of this approach is Noam Chomsky, who coined the term 'language organ' and argued in favour of the uniqueness and the innateness of human linguistic skills [40]. Most the biologists agree with Chomsky on this issue [2,37,41–43]. The empiricist paradigm, however, argues that linguistic performance by humans can be explained with domain-general learning techniques [44].

Language evolution
Non-adaptationist accounts of language evolution rely heavily on so-called 'spandrels' [45]. The idea is that language or linguistic skills evolved not because it gave a fitness advantage to its users, but because it was a side effect of other skills (some argue that recursion was used in some other domain, such as social cognition, and was taken over by language later [46]), as spandrels are a side effect of architectural constraints. Chomsky again has a prominent role in this debate as the protagonist of the non-adaptationist approach. In the latest reworking of the theory [46], Chomsky and colleagues distinguish between the so-called 'FLB' and 'FLN' of language (see Glossary). They argue that FLB consists of skills that evolved in other animals as well as in humans, whereas FLN consists of only one skill (merge), which evolved in a different (unspecified) context and was then co-opted for linguistic use. However, that European starlings appear able to recognise context-free grammatical structures (i.e. hierarchical syntax [47]) is somewhat contrary to Chomsky's position, given that it shows that a precursor (i.e. recognition of hierarchical, phrase-structure grammar) of the skill they have assigned to FLN (recognition and production of hierarchical syntactical structures through the 'merge' operation) has evolved independently in other animals too.

The first adaptationist account of human language was by Darwin [48], later defended by Pinker and Bloom [42] in their influential paper about the darwinian account of language. More specifically, these authors argued that language, as any complex adaptations, can only be explained by means of natural selection. This paper catalysed many linguists and biologists to study language and language evolution from the perspective of evolutionary biology and was followed by many influential publications [2,37,49,50]. Most recently, Jackendoff and Pinker [43] made a forceful defence of the adaptationist paradigm in response to Chomsky and colleagues [46].

suggests that a unique context and selective pressure were responsible for the evolution of human language, given that no other primates have yet evolved a language-like communication system despite the fact that, arguably, they have the basic cognitive skills required and a similar genetic background to humans.

In the light of this conclusion, it is interesting to compare those theories that attempt to explain the emergence of early human language. What can they say about the social context, the content, the first words and the selective advantage of early human language? Can they explain the uniqueness of human language in this (or any other) way?

The context of language evolution

Most of the theories that suggest a given context for the evolution of human language attempt to account for its functional role (Box 3). Given that, functionally, all of these theories are more or less plausible, it is almost impossible to decide on their usefulness based only on this criterion. However, recent game theoretical research can help us to evaluate various contexts. These criteria concern the interest of communicating parties and the cost of equilibrium signals.

The central issue is whether early linguistic communication was honest. If signal cost is the same for all signallers, then honest cost-free signalling can be

Box 2. The genetic background of human language

When dealing with any evolutionary transition, it is useful to know as much about the underlying genetic conditions as possible. In humans, twins provide a useful means for investigating the genetic background of various features, including language.

By surveying studies on the genetic background of language, Stromswold [51] concluded that concordance rates for language, including written (e.g. dyslexia) and spoken language [specific language impairment (SLI)] disorders are significantly higher for monozygotic than for dizygotic twins. This suggests that genetic factors have a role in such disorders. Stromswold concluded in her review and metaanalyis of the heritability of language that different genes might be responsible for the variance in different components of language (such as the rate of language acquisition in children and language proficiency in children and adults) and that some genetic effects might be specific to aspects of language (e.g. lexical versus syntactical abilities). The sum of all genetic effects is usually not greater than 50% of the total variance for various aspects of cognition (including language) [51] and most individual genes are expected to have small effects. However, because the cognitive skills of the sets of twins studied were not well described, intermediate phenotypes with known genetic backgrounds are required before firm conclusions can be made as to the involvement of genes in language. The situation is similar to that of geotaxis in *Drosophila,* where the individual involvement of different genes that collectively determine this capacity is counterintuitive [52]. For example, it is unlikely that anyone would have thought that the 'pigment dispersing factor' protein is involved in this ability (and, incidentally, also in circadian rhythms).

One example of a clear link between a gene and a language disorder was revealed in studies by Myrna Gopnik [53,54]. Gopnik described 'feature-blind' dysphasia (i.e. impaired marking in English of past tense and plural, for example) and found that its obvious genetic background was due to a single dominant allele of the gene, later identified as encoding the regulatory protein *FOXP2*. It is known that this gene has previously been under positive selection in the hominid lineage [55], which is consistent with the view that language is a genetically conditioned adaptation.

Researchers have since shown that, in humans and songbirds, both *FoxP2* and *FoxP1* are expressed in functionally similar brain regions that are involved in sensorimotor integration and skilled motor control [56]. Moreover, the differential expression of *FoxP2* in avian vocal learners is correlated with vocal plasticity [57]. In mice, which, similar to humans, also have two copies of the *Foxp2* gene, the ultrasonic vocalisation of pups separated from their mothers is severely affected if one copy of *Foxp2* is altered [58]. This suggests that this gene is involved in social communication across different species. However, further work is needed to elucidate more clearly the role of genes in language.

Box 3. Alternatives theories to explain language evolution

Gossip: menstrual ritual can be a costly signal of commitment; hence participating in such rituals can create female groups of shared interest in which sharing information about the social life of others (i.e. gossiping) can be beneficial [20].

Grooming hypothesis: language evolved as a substitution for physical grooming [25]. The need for this substitution derived from the increasing size of the early hominid groups, which mean that physical grooming became more time consuming, whereas it was possible to 'groom' more than one individual simultaneously via vocal communication.

Group bonding and/or ritual: language evolved in the context of intergroup rituals, which first occurred as a kind of 'strike action' against non-provisioning males. Once such rituals were established, a 'safe' environment was created for further language evolution [27].

Hunting theories: 'our intellect, interests, emotions, and basic social life – all are evolutionary products of the success of the hunting adaptation.' [59]. Later, Hewes in his paper about the gestural origins of language [22] takes up the idea and argues that the probable first use of language was to coordinate the hunting effort of the group.

Language as a mental tool: language evolved primarily for the function of thinking and was only later co-opted for the purpose of communication [30].

Mating contract and/or pair bonding: the increasing size of the early hominid groups and the need for male provisioning also necessitated 'social contract' between males and females [23].

Motherese: language evolved in the context of mother–child communication. Mothers had to put down their babies to collect food efficiently, and their only option to calm down babies was to use some form of vocal communication [24].

Sexual selection: language is a costly ornament that enables females to assess the fitness of a male. According to this theory, language is more elaborate than a pure survival function would require [28].

Song hypothesis: language evolved rapidly and only recently by a process of cultural evolution. The theory assumes two important sets of preadaptations; one is the ability to sing; the other is better representation abilities (i.e. thinking and mental syntax) [29].

Status for information: language evolved in the context of a so-called 'asymmetric cooperation', where information (that was beneficial to the group) was traded for status [21].

Tool making: assumes a double homology: 'a homologous neural substrate for early ontogeny of the hierarchical organisations shared by two domains – language and manual object combination – and a homologous neural substrate and behavioural organisation shared by human and non-human primates in phylogeny.' [26]

evolutionarily stable only if there is no conflict of interest between the participants [13]. If the cost of signals varies with the quality of the signaller, then the situation is more complicated. In this case, it is possible to construct cost functions that give an arbitrarily low cost at equilibrium even if there is a conflict of interest [14–16] (Box 4). In case of human language, the most obvious way to construct such a cost function is to punish dishonest signallers [16]. However, this solution assumes that dishonest signallers can, on average, be detected (i.e. signals can be cross-checked); it also assumes that dishonest signallers are punished (which is a non-trivial assumption). Thus, one can conclude that, "'conventional' signals will be used when communicating about (i) coincident interest or (ii) verifiable aspects of conflicting interest; 'costly' signals will be used otherwise." [16] Although theory so far says nothing about the evolution of such systems of communication, there are a few computer simulations that suggest that honest cost-free communication evolves only if there is shared interest between the participants [17–19].

What does this tells us about the emergence of human language? The production cost of speech or gesturing appears to be low [4], thus human language consists of cost-free or low-cost signals at equilibrium (not counting time constraints). Thus, based on the above criteria, one should favour either those theories that propose a context with no conflict of interest (e.g. hunting, tool making, motherese, grooming or the group bonding and/or ritual theory) or a context in which there might be a conflict of interest but signals can be easily cross-checked. None of the theories fit the second context: for example, mating contract and gossiping both assume a context in which conflict of interest exists and signals cannot be easily cross-checked.

What was communicated?

Even in a given context, it is not obvious what was communicated, especially if there was conflict of interest between participants. Some theories make explicit statements about the assumed information content of early

Box 4. The problem of honesty

There is an ongoing debate about the honesty of animal communication, centred around the proposition that signals need to be costly to be honest [60]. Although some models appear to provide support for this statement [13,61,62], there are exceptions. First, cost-free signals can be evolutionarily stable provided that there is no conflict of interest between the communicating parties [13]. Second, even where a conflict of interest exists, cost-free signals can be evolutionarily stable provided that the cost of signals is a function of the quality being signalled [14–16]. The most general case is when the fitness of the signaller depends on its state and the fitness of both players is influenced by the survival of the other. Assuming a discrete model with two states, two signals and two responses, the conditions of evolutionarily stable honest signalling are as follows (Equations I–VI) [15]:

$$W_h + rV_h > 0 \quad \text{[EqnI]}$$

$$W_l + rV_l < 0 \quad \text{[EqnII]}$$

$$V_h + rW_h > C_h \quad \text{[EqnIII]}$$

$$V_l + rW_l < C_l \quad \text{[EqnIV]}$$

$$V_h + rW_h > 0 \quad \text{[EqnV]}$$

$$V_l + rW_l > 0 \quad \text{[EqnVI]}$$

where *W*, *V* and *C* denote the fitness of the receiver and signaller, and the cost of signalling, respectively. *l* and *h* denote the quality of the signaller ('high' and 'low', respectively). The fitness of each player can be influenced by the survival of the other player (*r*). Equation I and II describe the conditions of the receiver for honest signalling; Equations III and IV reflect are the condition of the signaller; and Equations V and VI describe conflicts of interest. Reversing the inequality in Equation VI would mean that there is no conflict of interest between signaller and receiver. If $r = 1$, then there can be no conflict of interest, assuming that signalling is beneficial for the receiver (because $V_l + rW_l = W_l + rV_l$, Equation VI cannot be fulfilled). This implies that, in this case, C_l can equally zero or less than zero (given that the left-hand side of Equation IV need not be greater than zero). However, if $0 < r < 1$ then at least C_l should be greater than zero (Equation IV). That is, signalling for low-quality individuals must be costly in case of a conflict of interest. Signalling for high-quality individuals need not be costly even in this case (Equation III). Given that at the honest equilibrium only high-quality individuals signal, the observed cost can be zero. If, however, $C_l = C_h$ then both costs should be greater than zero for honest signalling to be stable. Figure I depicts the regions of honest signalling in case of conflict of interest. The same logic also applies and, thus, the same results hold, for continuous models [16].

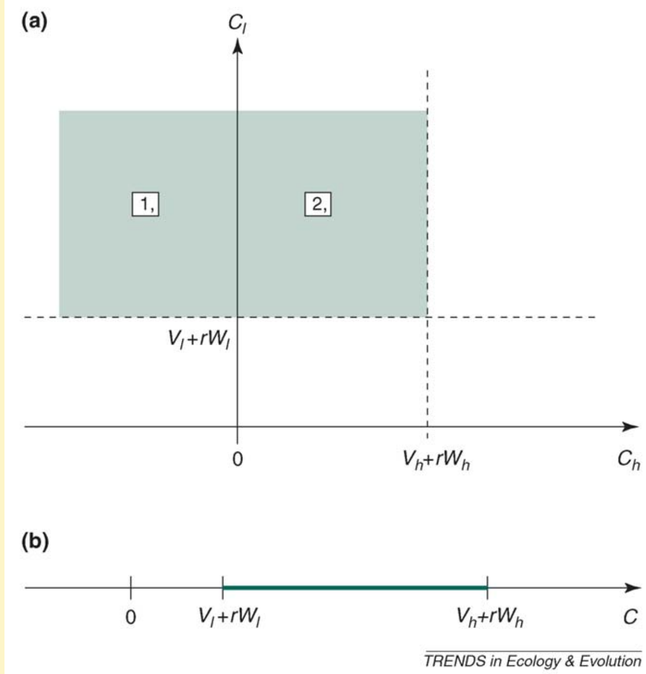

Figure I. The relationship between signal cost and evolutionarily stable honest signalling. **(a)** Regions of honest evolutionarily stable signalling: 1st, 2nd (green shading) when $C_h \neq C_l$, $V_h \neq V_l$ and $r > 0$. C_l must be greater than zero, whereas C_h can equal zero or even less than zero. **(b)** Regions of honest evolutionarily stable signalling when $C_h = C_l$ (thick line); both C_h and C_l must be greater than zero. In cases of shared interest ($r = 1$), both C_h and C_l can equal zero or even less than zero (see Equations IV, VI) [15].

linguistic communication, whereas others leave us in the dark. The suggested information content can be classified as: (i) socially important information about the state and behaviour of group members (i.e. gossip, status for information, [20,21]); (ii) information about the relevant states of the environment (i.e. status for information, hunting theories, [21,22]); (iii) future behaviour (i.e. mating contract, [23]); (iv) contact calls, information about the proximity of group members (i.e. motherese, [24]); (v) social bonding with no information content (grooming, [25]) and (vi) tool making [26]. In addition, there are theories that give no hint about the content of communication. These include mental tool, sexual selection, song theory and group bonding and/or ritual [27–30].

What were the first words?

Even in case of those theories that hint as to the content of communication, only a few hint about the possible first words. Although this appears to be a difficult task, several attempts were recently made to model the evolution of meaning using word pairs [31]. In these experiments, (virtual) agents play the so-called 'naming' (or guessing) game, during which word–meaning pairs evolve autonomously in a population of agents. The game itself is simple in that there are two agents: one is the speaker, the other the listener. The speaker 'thinks' of an object out of a set of available objects (this set is known to the listener) and gives it a name that gives the highest chance, according to the speaker, that the listener can find out which object was picked. If the listener 'understands' successfully then this word–meaning pair is reinforced in both agents; if not, then the link is weakened.

Steels and colleagues [31] identified the following conditions that the agents should fulfil for the game to succeed: (i) agents must be able to engage in coordinated interactions; (ii) agents must have parallel non-verbal ways to achieve the goals of interactions (e.g. pointing); (iii) agents must have ways to conceptualise reality and to acquire these conceptualisations, constrained by the semantic concepts expressed in the emerging lexicon and the types of situations they encounter; (iv) agents must have ways to recognise word forms and reproduce them; and (v) agents must have the ability to discover and use the strongest associations (between words and meanings) in the group. The second and third points are especially noteworthy given that they imply that abstract concepts (which cannot be pointed at and are not easy to conceptualise) are unlikely to be the first words. Pointing also requires that the agents can see each other and that the referent should be visible to both participants.

Most of the theories do not consider groundedness and do not say anything about the possible first words; a handful propose highly abstract words as the first tokens, such as 'faithful' or 'philander' (gossip [20]) and 'faithfulness' (mating contract [23]). Only the motherese and the hunting theories [22,24] suggest words ('mama' and names of prey animals, respectively) that can be easily grounded in reality.

The power of generalisation

One defining feature of human language is the range and power of generalisations it allows. The topic of conversation is not linked to the observable present; humans can talk about the past, present and future, and about distant places; they can make generalisations about living and non-living; can invent persons, animals and stories that never existed. Can any of the theories explain this capacity? Perhaps it is a little unfair to expect this from these theories, but one can ask whether any of them assume a level of generalisation above that of the self-reporting signals used by animals.

The result is mixed. Some of the theories, such as contact calls or grooming [24,25], assume contexts in which no reasonable generalisations are expected to be made. Some (mental tool, song theory [29,30]) do not specify the context but assume mental syntax as a pre-requisite for language, thus assuming some form of already existing generalisation capacity. Other theories, such as gossip, mating contract, tool making and hunting, propose contexts in which generalisations can be made on observable events (e.g. the type of observable behaviours, or type of prey animals). Last but not least, some theories (status for information, sexual selection) assume this capacity yet do not give an exact context of these generalisations.

Assuming that our generalisation capacity is selected for (i.e. it gave selective advantage to those individuals that were able to make and understand generalisations), theories should be favoured that suggest an exact context for these generalisations and in which these generalisations can be made on observable events.

Selective forces behind language evolution: uniqueness

Given the uniqueness of human language, one should be able to explain why other species living under similar conditions did not evolve language, a fact that strongly suggests that human language is a special adaptation. Hurford [32] argues in favour of a similar approach: '…in general, more realistically and more eclectically, for any set of circumstances proposed as individually necessary and collectively sufficient to explain the emergence of Language, one has to show that this combination of circumstances applies (or applied) to humans and to no other species.'

Most of the theories do not consider the kind of selective forces that could encourage the use of conventional communication in a given context instead of the use of 'traditional' animal signals. Most contexts put forth to explain language evolution can be found in animals: mate choice, pair bonding, contact calls, parent–offspring communication all feature prominently in the lives of many animal species [33] yet none evolved a communication system comparable to human language. For example, chimpanzees have a fascinating and complex social life [34], but have not yet evolved a complex system of symbolic communication, although, arguably, they would benefit from gossiping. There is a rich tradition of alliance making and/or breaking, peace making [34], in general the kind of social life that some would propose as the cradle of language evolution [20]. There are also group-living mammalian predators (e.g. African wild dogs) that can carry out coordinated hunting without any system of symbolic communication. Finally, the primary skill for reliable tool making is precise imitation, which is found in humans [35]. Thus, there is no theory that convincingly demonstrates a situation that would require a complex means of symbolic communication rather than the existing simpler communication systems.

Conclusions

Explaining the evolution of human language is likely to remain a challenge for the coming decade. As we have discussed, there is no single theory that could sufficiently answer all the questions about honesty and groundedness, power of generalisation, and uniqueness. Table 1 summarises these criteria. As one can see, most of the theories fail to answer most of the questions. Perhaps the easiest criterion to fulfil is shared interest, as there are

Table 1. The properties and explanatory powers of competing language evolution theories

Hypothesis	Modality	First words[b]	Topic[b]	Question[a] (i)	(ii)	(iii)	(iv)	Refs
Gossip	V	'Faithful', 'Philander'	Social life	No	No	Yes	No	[20]
Grooming hypothesis	V	?	?	Yes	No	No	No	[25]
Group bonding and/or ritual	?/V	?	?	Yes	No	No	No	[27]
Hunting theories	G/V	Prey animals	Coordination of the hunt	Yes	Yes	Yes	No	[22,59]
Language as a mental tool	T	?	?	Yes	No	Yes	No	[30]
Mating contract and/or pair bonding	?	?	Social contract	No	No	No	No	[23]
Motherese	V	'Mama'	Contact call	Yes	Yes	No	No	[24]
Sexual selection	?	?	Anything	No	No	No	No	[28]
Song hypothesis	V	?	?	No	No	No	No	[29]
Status for information	?	?	Valuable information	No	No	Yes	No	[21]
Tool making	?	?	?	Yes	Yes	Yes	No	[26]

[a]The following questions were asked in order to evaluate the alternative theories: (i) Honesty: can the theory account for the honesty of early language, that is, is there a shared interest between the proposed communicating parties? (ii) Groundedness: are the concepts proposed by the theory grounded in reality? (iii) Power of generalisation: can the theory account for the power of generalisation, which is unique to human language? (iv) Uniqueness: can the theory account for the uniqueness of human language? Most of the theories can answer only one or two questions, some none at all; only the tool making and hunting theories can answer three questions out of four. The most notable weakness of the theories is that none of them can explain convincingly the uniqueness of human language. Thus, it remains a challenge not just to propose a scenario that can answer all four questions, but also to explain the uniqueness of human language in the first place.
[b]'?': no information available; G, gestures; T, thought, V: vocalisation.

several social situations that assume shared interest between communicating parties (such as hunting or contact calls). There are only two theories, 'tool making' and 'hunting' [22,26], that do significantly better than the others as they can answer three out of the four questions asked of them (Table 1). Thus, it might be tempting to say that some combination of the two could provide a series of selective scenarios that would fit all of our criteria. The most notable conclusion, however, is that all the theories fail to explain the uniqueness of human language. Thus, even though indirect evidence strongly suggests that the evolution of human language was selection limited, it remains difficult to envisage a scenario that would show why.

Although the different scenarios suggest all kinds of selective forces, none of these scenarios has been consistently implemented in a family of models. Given the limitations on experimentation on humans and chimps, researchers should consider implementing the different scenarios in various model-based settings. Ultimately, researchers should be able to re-enact the emergence of language in artificial worlds, many of which will probably involve robots (e.g. http://ecagents.istc.cnr.it/). The use of robots offers a unique and probably indispensable way of symbol grounding (i.e. basic words, via concepts, should be linked to physical reality) and somatosensory feedback (i.e. actions, or results of actions, on behalf of the agent, feed back into its own cognitive system via sensory channels [36].

Such ambitious research projects would have associated difficulties too numerous to mention here; however, the tasks that the agents (e.g. robots) would be subjected to imply a complicated fitness landscape that is similar to climbing a staircase rather than a hill: a good capacity for imitation would probably coevolve with a capacity to learn symbols (words), which then opens up the possibility to climb to the first level ('stair') of syntax [37]. Successful modelling could answer the burning question of how genes, under the influence of natural selection, could rig a neuronal system so that it becomes able to handle linguistic input and output at the level of symbolic reference combined with complex syntax [38].

The only process that appears to have solved the 'language problem' is evolution by natural selection. But there is no guarantee that just any kind of selection scenario, even if implemented *in silico,* would lead to the origin of such a faculty, partly owing to the results of the analysis presented here, partly because of what is known as the 'no free lunches theorem' [39], which states that the efficiency of an evolutionary search process is dependent on the problem. Putting constraints on the selective scenarios might constrain the search space to such an extent that simulated evolution will be able to re-enact the fascinating evolutionary transition of language evolution.

Acknowledgements

This work was partly supported by the Hungarian National Research Fund (OTKA T047245), National Office for Research and Technology (NAP 2005/KCKHA005) and the ECAGENTS project (http://ecagents.istc.cnr.it/). The ECAGENTS project is funded by the Future and Emerging Technologies Programme (IST-FET) of the European Community under EU R&D contract IST-2003-1940. The information provided is the sole responsibility of the authors and does not reflect the Community's opinion. The Community is not responsible for any use that might be made of data appearing in this publication.

References

1 Christiansen, M. and Kirby, S. (2003) Language evolution: the hardest problem in science? In *Language Evolution* (Christiansen, M. and Kirby, S., eds), pp. 1–15, Oxford University Press
2 Maynard Smith, J. and Szathmáry, E. (1995) *Major Transitions in Evolution,* Freeman
3 Jablonka, E. and Szathmáry, E. (1995) The evolution of information storage and heredity. *Trends Ecol. Evol.* 10, 206–211
4 Maynard Smith, J. and Harper, D. (2003) *Animal Signals,* Oxford University Press
5 Fehr, E. and Fischbacher, U. (2004) Social norms and human cooperation. *Trends Cogn. Sci.* 8, 185–190
6 Avital, E. and Jablonka, E. (2000) *Animal Traditions: Behavioural Inheritance in Evolution,* Cambridge University Press
7 Szathmáry, E. (2003) Cultural processes: the latest major transition in evolution. In *Encyclopedia of Cognitive Science* (Nadel, L., ed.), Macmillan, DOI: 10.1002/0470018860.s00716
8 Cavalier-Smith, T. (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int. J. Syst. Evol. Microbiol.* 52, 297–354
9 Savage-Rumbaugh, S. *et al.* (1998) *Apes, Language and the Human Mind,* Oxford University Press
10 The Chimpanzee Sequencing and Analysis Consortium (2005) Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* 437, 69–87
11 Szathmáry, E. (2001) Origin of the human language faculty: the language amoeba hypothesis. In *New Essays on the Origin of Language* (Trabant, J. and Ward, S., eds), pp. 41–51, Mouton/de Gruyter
12 Conway Morris, S. (2003) *Life's Solution: Inevitable Humans in a Lonely Universe,* Cambridge University Press
13 Maynard Smith, J. (1991) Honest signalling: the Philip Sidney game. *Anim. Behav.* 42, 1034–1035
14 Hurd, P.L. (1995) Communication in discrete action-response games. *J. Theor. Biol.* 174, 217–222
15 Számadó, Sz. (1999) The validity of the handicap principle in discrete action-response games. *J. Theor. Biol.* 198, 593–602
16 Lachmann, M. *et al.* (2001) Cost and constraints in animals and in human language. *Proc. Natl. Acad. Sci. U. S. A.* 28, 13189–13194
17 Bullock, S. (1998) A continuous evolutionary simulation model of the attainability of honest signalling equilibria. In *Artificial Life VI.: Proceedings of the Sixth International Conference on Artificial Life* (Adami, A. *et al.*, eds), pp. 339–348, MIT Press/Bradford Books
18 Noble, J. (2000) Cooperation, competition and the evolution of prelinguistic communication. In *Evolutionary Emergence of Language* (Knight, C. *et al.*, eds), pp. 40–61, Cambridge University Press
19 Harris, D. and Bullock, S. (2002) Enhancing game theory with coevolutionary simulation models of honest signalling. In *Congress on Evolutionary Computation* (Fogel, D., ed.), pp. 1594–1599, IEEE Press
20 Power, C. (1998) Old wives' tales: the gossip hypothesis and the reliability of cheap signals. In *Approaches to the Evolution of Language* (Hurford, J.R. *et al.*, eds), pp. 111–129, Cambridge University Press
21 Dessalles, J. (1998) Altruism, status and the origin of relevance. In *Approaches to the Evolution of Language* (Hurford, J.R. *et al.*, eds), pp. 130–147, Cambridge University Press
22 Hewes, G. (1973) Primate communication and the gestural origin of language. *Curr. Anthropol.* 14, 5–25
23 Deacon, T. (1997) *The Symbolic Species,* Penguin Books
24 Falk, D. (2004) Prelinguistic evolution in early hominins: whence motherese? *Behav Brain Sci.* 27, 491–503
25 Dunbar, R. (1998) Theory of mind and the evolution of language. In *Approaches to the Evolution of Language* (Hurford, J.R. *et al.*, eds), pp. 92–110, Cambridge University Press
26 Greenfield, P.M. (1991) Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behaviour. *Behav. Brain Sci.* 14, 531–595
27 Knight, C. (1998) Ritual/speech coevolution: a solution to the problem of deception. In *Approaches to the Evolution of Language* (Hurford, J.R. *et al.*, eds), pp. 68–91, Cambridge University Press

28 Miller, G. (2001) *The Mating Mind,* Anchor Books
29 Vaneechoutte, M. and Skoyles, J.R. (1998) The memetic origin of language: modern humans as musical primates. *J. Memet.* 2, (http://jom-emit.cfpm.org) /1998/vol2/vaneechoutte_m&skoyles_jr.html
30 Burling, R. (1993) Primate calls, human language, and nonverbal communication. *Curr. Anthropol.* 34, 25–53
31 Steels, L. *et al.* (2002) Crucial factors in the origin of word-meaning. In *The Transition to Language* (Wray, A., ed.), pp. 252–271, Oxford University Press
32 Hurford, J.R. (1999) The evolution of language and languages. In *The Evolution of Culture* (Dunbar, R. *et al.*, eds), pp. 173–193, Edinburgh University Press
33 Hauser, M.D. (1996) *The Evolution of Communication,* MIT Press
34 de Waal, F. (1982) *Chimpanzee Politics,* Harper & Row
35 Nagel, K. *et al.* (1993) Processes of social learning in the tools use of chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *J. Comp. Psychol.* 107, 174–186
36 Nolfi, S. and Floreano, D. (2002) Synthesis of autonomous robots through evolution. *Trends Cogn. Sci.* 6, 31–37
37 Jackendoff, R.S. (1992) *Languages of the Mind,* MIT Press
38 Szathmáry, E. *et al*. *In silico* evolutionary developmental neurobiology and the origin of natural language. In *Emergence of Communication and Language* (Lyon, C. *et al.*, eds), Springer-Verlag (in press)
39 Wolpert, D.H. and Macready, W.G. (1997) No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* 1, 67–82
40 Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin and Use,* Praeger
41 Pinker, S. (1994) *The Language Instinct,* Penguin Books
42 Pinker, S. and Bloom, P. (1990) Natural language and natural selection. *Behav. Brain Sci.* 13, 707–786
43 Pinker, S. and Jackendoff, R. (2005) The faculty of language: what's special about it? *Cognition* 95, 201–236
44 Sampson, G. (1997) *Educating Eve: The 'Language Instinct' Debate,* Cassell
45 Gould, S.J. and Lewontin, R.C. (1979) The Spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proc. R. Soc. B* 205, 581–598
46 Hauser, M.D. *et al.* (2002) The faculty of language: what is it, who has it, and how did it evolve? *Science* 298, 1569–1579
47 Gentner, T.Q. *et al.* (2006) Recursive syntactic learning by songbirds. *Nature* 440, 1204–1207
48 Darwin, C.R. (1871) *The Descent of Man and Selection in Relation to Sex,* John Murray
49 Christiansen, M. and Kirby, S., (eds) (2003) *Language Evolution*, Oxford University Press
50 Knight, C. *et al.* (eds) (2000) *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, Cambridge University Press
51 Stromswold, K. (2001) The heritability of language: a review and metaanalysis of twin, adoption, and linkage studies. *Language* 77, 647–723
52 Toma, D.P. *et al.* (2002) Identification of genes involved in *Drosophila* geotaxis, a complex behavioural trait. *Nat. Genet.* 31, 349–353
53 Gopnik, M. (1990) Feature-blind grammar and dysphasia. *Nature* 344, 715
54 Gopnik, M. (1999) Familial language impairment: more English evidence. *Folia Phoniatr. Logop.* 51, 5–19
55 Enard, W. *et al.* (2002) Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418, 869–872
56 Teramitsu, I. *et al.* (2004) Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *J. Neurosci.* 24, 3152–3163
57 Haesler, S. *et al.* (2004) FoxP2 expression in avian vocal learners and non-learners. *J. Neurosci.* 24, 3164–3175
58 Shu, W. *et al.* (2004) Altered ultrasonic vocalization in mice with a disruption in the *Foxp2* gene. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9643–9648
59 Washburn, S.L. and Lancaster, C. (1968) The evolution of hunting. In *Man the Hunter* (Lee, R.B. and DeVore, I., eds), pp. 293–303, Aldine
60 Zahavi, A. (1975) Mate selection – a selection for handicap. *J. Theor. Biol.* 53, 205–214
61 Grafen, A. (1990) Biological signals as handicaps. *J. Theor. Biol.* 144, 517–546
62 Godfray, H.C.J. (1991) Signalling of need by offspring to their parents. *Nature* 352, 328–330