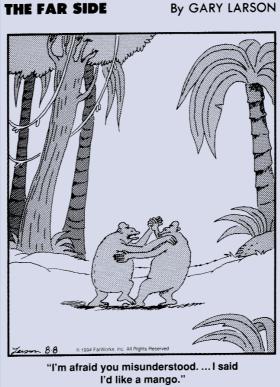# Survival of the clearest

Steven Pinker

There are no fossils to show how language evolved. But evolutionary game theory is revealing how some of the defining features of human language could have been shaped by natural selection.

The study of the evolution of language, famously banned by the Société de Linguistique de Paris in 1866 and dismissed as idle story-telling ever since, has returned to respectability. Recent years have seen a flurry of articles and books[1–5], a biannual research conference and now several papers by Martin Nowak and collaborators applying evolutionary game theory to the problem[6–9]. Their latest offering appears on page 495 of this issue[9].

Language would seem to be a natural target of evolutionary analysis. Universal, heritable, rapidly acquired by children and well suited to communicating complex information, language is plausibly an adaptation that allowed our ancestors to share knowledge and negotiate agreements, fuelling the technological and demographic explosions that led our species to infest the planet millennia ago.

But the very idea of evolutionary linguistics, like evolutionary psychology in general, draws mixed reactions from scientists. On the one hand, "nothing in biology makes sense except in the light of evolution." The major parts of our psychology could no more be understood without considering their biological function than could the heart or eye. For example, visual perception does not just entertain us with pretty colours and shapes, but delivers usable information about the surfaces and materials of the world. And it is no surprise that people everywhere fear heights and snakes more than comfy chairs and bunnies. On the other hand, other evolutionary 'explanations' of behaviour — "the function of humour is to relieve tension", "music evolved to foster group cohesion" — can seem glib and lame.

Fortunately, good theories of adaptation can be distinguished from bad ones[10]. The bad ones try to explain one bit of our psychology (say, humour or music) by appealing to some other, equally mysterious bit (laughing makes you feel better; people like to make music with other people). The good ones use some independently established finding of science or mathematics to show that some mechanism can attain some goal in some environment. For instance, projective geometry shows how stereoscopic imaging can compute the third dimension; herpetology has documented that many snakes are venomous. These engineering bench-

THE FAR SIDE By GARY LARSON

"I'm afraid you misunderstood. ... I said I'd like a mango."

Figure 1 **According to recent work in evolutionary game theory — the latest example of which appears elsewhere in this issue[9] — minimizing communication errors was a major selection pressure in the evolution of human language.**

marks can serve as predictions for how Darwinian organisms ought to work. The more uncannily the engineering specifications for a system match the facts of the organism, the more confidently we can infer that the organism was selected to perform that function.

With many behavioural systems, the relevant environment consists not of predictable things like light waves and snakes but of other behaving individuals co-evolving their own strategies. Evolutionary game theory[11] has allowed biologists to analyse such dynamics, predicting how Darwinian organisms ought to treat their antagonists, mates and comrades.

Language, like sex, aggression and cooperation, is a game it takes two to play. Many origins have been proposed for it — a substitute for grooming, a way to tattle on adulterers, a by-product of a big head — but they have been based on intuitive appeal. Game theory can provide the external criteria for utility enjoyed by the rest of evolutionary biology. Modellers make the minimal assumption that the transmission of infor-

mation between partners provides them with an advantage (say, by exchanging know-how or coordinating their behaviour), and that the advantage translates into more offspring, with similar communicative skills. The question then is how a stable communication system might evolve from repeated pairwise interactions.

In 1989 the linguist James Hurford[12] used evolutionary game theory to show that a defining property of human language, the arbitrary, bidirectional sign (in which a community of speakers and hearers tacitly agree to use particular signals to refer to particular concepts), will drive out other schemes over evolutionary time. Nowak and collaborators have done the same for two other defining properties of language.

Last year, Nowak and David Krakauer[6] pointed out that real organisms never transmit information perfectly: errors in signalling or perception are inevitable, especially when signals are physically similar (Fig. 1). Imagine organisms that use a different sound (say, a vowel) for every concept they wish to communicate. As they commu-

441

nicate more concepts, they will need additional sounds, which will be physically closer and hence harder to discriminate. At some point adding new signals just makes the whole repertoire more confusable and fails to increase its net communicative power.

Nowak and Krakauer showed that this limitation can be overcome by capping the number of signals and stringing them together into sequences, with one sequence per concept. The sequences are what we call words, and the combination of meaningless vowels and consonants into meaningful words (d + o + g = dog) by rules of phonology is a universal property of language: half of what linguists call 'duality of patterning'. Nowak and Krakauer demonstrated that its evolution is likely among communicators with a large number of messages to convey, a precondition that plausibly characterizes our brainy ancestors.

Now, Nowak, Joshua Plotkin and Vincent Jansen[9] analyse another hallmark. For a word to survive in a community, it must be used frequently enough to be heard and remembered by all the learners. As new words are added to the vocabularies of speakers, old words must be used less often, and they are liable to fade, leaving the language no more expressive than before.

Nowak et al. point out that this limitation can be overcome by communicators who use compositional syntax: rather than pairing each word with an entire event, they pair each word with a component of an event (a participant, an action, a relationship), and string the words together in an order that reflects their roles (for example, dog bites man). Such communicators need not memorize a word for every event, reducing the word-learning burden and allowing them to talk about events that lack words. Syntax and semantics, the other half of the duality of patterning, will evolve.

Nowak et al.[9] note that syntax has a cost: the requirement to attend to the order of words. Its benefits exceed the costs only when the number of events worth communicating exceeds a threshold. This 'syntax threshold' is most likely to be crossed when the environment, as perceived by the communicators, has a combinatorial structure, for example, when any of a number of actors (dogs, cats, men, women, children) can engage in any of a number of actions (walking, running, sleeping, biting)[1]. In such a world, the number of words that have to be learned by a syntactic communicator equals the sum of the number of actors, actions, places and so on, whereas the number that must be learned by a nonsyntactic communicator equals their product, a potentially unlearnable number. Syntax is handy to an analytical mind in a combinatorial world.

Of course, showing how language could have evolved in a hypothetical world is a far cry from showing how it did evolve in this world. But the game theorists have demonstrated the evolvability of the most striking features of language — arbitrary signs and duality of patterning — breaking the logical circle that cripples many evolutionary explanations. And by showing how the abstract components of language (lexicon, phonology, syntax and semantics) might be related to more tangible skills, such as the articulation and perception of speech and the cognitive categorization of the world, they suggest ways of connecting the evolution of language to other topics in human evolution, allowing each to constrain the others. Perhaps some day the Société will reverse its ban. ■

Steven Pinker is in the Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02193, USA.
e-mail: steve@psyche.mit.edu

1. Pinker, S. & Bloom, P. *Behav. Brain Sci.* **13,** 707–784 (1992).
2. Pinker, S. *The Language Instinct* (HarperCollins, New York, 1994).
3. Hurford, J. R., Studdert-Kennedy, M. & Knight, C. (eds) *Approaches to the Evolution of Language: Social and Cognitive Bases* (Cambridge Univ. Press, New York, 1998).
4. Deacon, T. W. *The Symbolic Species: The Co-evolution of Language and the Brain* (Norton, New York, 1997).
5. Calvin, W. H. & Bickerton, D. *Lingua ex Machina: Reconciling Darwin and Chomsky with the Human Brain* (MIT Press, Cambridge, Massachusetts, 2000).
6. Nowak, M. A. & Krakauer, D. C. *Proc. Natl Acad. Sci. USA* **96,** 8028–8033 (1999).
7. Nowak, M. A., Krakauer, D. C. & Dress, A. *Proc. R. Soc. Lond. B* **266,** 2131–2136 (1999).
8. Nowak, M. A., Plotkin, J. B. & Krakauer, D. C. *J. Theor. Biol.* **200,** 147–162 (1999).
9. Nowak, M. A., Plotkin, J. B. & Jansen, V. A. A. *Nature* **404,** 495–498 (2000).
10. Williams, G. C. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought* (Princeton Univ. Press, 1966).
11. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge Univ. Press, New York, 1982).
12. Hurford, J. R. *Lingua* **77,** 187–222 (1989).

Genome sequencing

A view of Mount *Drosophila*

Jonathan Hodgkin

Biologists live in privileged times. In the next few years, the complete genome sequences of representatives from all the major groups of organisms on Earth will be determined. Like a range of mountains emerging from cloud, successively larger and larger genomes can now be seen in their entirety, as each major genome project reaches its goals. Every one of these projects creates a gigantic and breathtaking archive of solid data and new information, which can be analysed in detail for decades to come. Nothing like this has happened in previous scientific history, or is likely to happen again.

The latest grand achievement in genomics has a particular significance for the science of genetics. Almost a century of research on the fruitfly *Drosophila melanogaster* (Fig. 1) has culminated in the publication of 120 million base pairs (Mb) of sequence[1], and it seems to include the entire protein-coding part of the fly genome. Although a number of bacterial (prokaryotic) genomes have been sequenced, this is only the third eukaryotic genome to arrive at this level of coverage. It follows the 12 Mb of the yeast *Saccharomyces cerevisiae*[2] and the 97 Mb of the nematode worm *Caenorhabditis elegans*[3], and is celebrated in a whole series of papers in *Science*[1,4–9]. These report both the technical aspects of the work and some of the first insights to emerge from a complete panorama of the genome.

Technical issues will not be discussed further here, but are especially important in view of the whole-genome shotgun strategy used for this project, in contrast to the clone-by-clone approach being used by the pub-

Figure 1 **In the limelight — the fruitfly** *Drosophila melanogaster.* **The sequence of the genome of** *Drosophila* **has now been published[1], giving geneticists invaluable new information about their favourite organism.**

licly funded human genome consortium. It is still unclear whether the shotgun strategy will work when applied to the much larger genomes of vertebrates and economically important plants[4].

The sequence determined does not correspond to the whole 180 Mb that make up the four chromosomes of the fly, and it still contains gaps and some stretches of low quality. The other 60 Mb (technically, heterochromatin) remains largely uncloned, unsequenced and neglected because most of it consists only of monotonous repeats and dead transposons.

As with most other animal and plant genomes, determining every last nucleotide in the sequence is not justified. The authors declare their sequence "substantially complete", in contrast to the "essentially complete" sequence of *C. elegans*, of which more