The evolution of language
Babel or babble?

Languages all have their roots in the same part of the world. But they are not as similar to each other as was once thought

Apr 14th 2011 | From the print edition

WHERE do languages come from? That is a question as old as human beings' ability to pose it. But it has two sorts of answer. The first is evolutionary: when and where human banter was first heard. The second is ontological: how an individual human acquires the power of speech and understanding. This week, by a neat coincidence, has seen the publication of papers addressing both of these conundrums.

Quentin Atkinson, of the University of Auckland, in New Zealand, has been looking at the evolutionary issue, trying to locate the birthplace of the first language. Michael Dunn, of the Max Planck Institute for Psycholinguistics in the Netherlands, has been examining ontology. Fittingly, they have published their results in the two greatest rivals of scientific journalism. Dr Atkinson's paper appears in *Science*, Dr Dunn's in *Nature*.

Travellers' tales

The obvious place to look for the evolutionary origin of language is the cradle of humanity, Africa. And, to cut a long story short, it is to Africa that Dr Atkinson does trace things. In doing so, he knocks on the head any lingering suggestion that language originated more than once.

One of the lines of evidence which show humanity's African origins is that the farther you get from that continent, the less diverse, genetically speaking, people are. Being descended from small groups of relatively recent migrants, they are more inbred than their African forebears.

Dr Atkinson wondered whether the same might be true of languages. To find out, he looked not at genes but at phonemes. These are the smallest sounds which differentiate meaning (like the "th" in thin; replace it with "f" or "s" and the result is a different word). It has been known for a while that the less widely spoken a language is, the fewer the phonemes it has. So, as groups of people ventured ever farther from their African homeland, their phonemic repertoires should have dwindled, just as their genetic ones did.

To check whether this is the case, Dr Atkinson took 504 languages and plotted the number of phonemes in each (corrected for recent population growth, when significant) against the distance between the place where the language is spoken and 2,500 putative points of origin, scattered across the world. The relationship that emerges suggests the actual point of origin is in central or southern Africa (see chart), and that all modern languages do, indeed, have a common root.

Lost in translation

Phonetic complexity and geographical location of languages

— Line of best fit for 504 languages*

Total phoneme diversity, standard deviations

Ewondo — Burmese — Bobo Fing — Cantonese — French — Kurdish — English — Bengali — Korean — Zulu — Russian — MEAN — Greek — Hawaiian — Japanese — Basque — Maori — Arabic (Egyptian) — Hebrew — Awa Pit

Distance from central Africa ('000km)

Source: *Science*   *Selected languages shown

That fits nicely with the idea that being able to speak and be spoken to is a specific adaptation—a virtual organ, if you like—that is humanity's killer app in the struggle for biological dominance. Once it arose, *Homo sapiens* really could go forth and multiply and fill the Earth.

The details of this virtual organ are the subject of Dr Dunn's paper. Confusingly, though, for this neat story of human imperialism, his result challenges the leading hypothesis about the nature of the language organ itself.

Grammar or just rhetoric?

The originator of that hypothesis is Noam Chomsky, a linguist at the Massachusetts Institute of Technology. Dr Chomsky argues that the human brain comes equipped with a hard-wired universal grammar—a language instinct, in the elegant phrase of his one-time colleague Steven Pinker. This would explain why children learn to speak almost effortlessly.

The problem with the idea of a language instinct is that languages differ not just in their vocabularies, which are learned, but in their grammatical rules, which are the sort of thing that might be expected to be instinctive. Dr Chomsky's response is that this diversity, like the
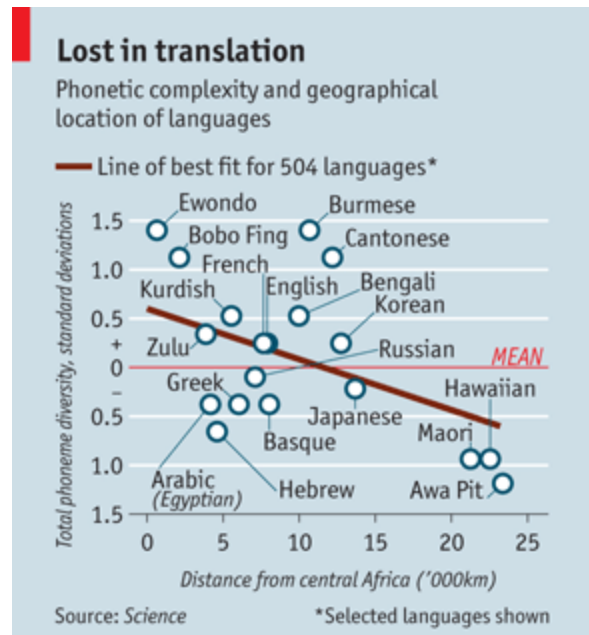
diversity of vocabulary, is superficial. In his opinion grammar is a collection of modules, each containing assorted features. Switching on a module activates all these features at a stroke. You cannot pick and choose within a module.

For instance, languages in which verbs precede objects will always have relative clauses after nouns; a language cannot have one but not the other. A lot of similar examples were collected by Joseph Greenberg, a linguist based at Stanford, who died in 2001. And, though Greenberg himself attributed his findings to general constraints on human thought rather than to language-specific switches in the brain, his findings also agree with the Chomskyan view of the world. Truly testing that view, though, is hard. The human brain cannot easily handle the connections that need to be made to do so. Dr Dunn therefore offered the task to a computer. And what he found surprised him.

Place your bets

To find out which linguistic features travel together, and might thus be parts of Chomskyan modules, means drawing up a reliable linguistic family tree. That is tricky. Unlike biologists, linguists do not have fossils to guide them through the past (apart from a few thousand years of records from the few tongues spoken by literate societies). Also, languages can crossbreed in a way that species do not. English, for example, is famously a muddle of German, Norse and medieval French. As a result, linguists often disagree about which tongues belong to a particular family.

To leap this hurdle, Dr Dunn began by collecting basic vocabulary terms—words for body parts, kinship, simple verbs and the like—for four large language families that all linguists agree are real. These are Indo-European, Bantu, Austronesian (from South-East Asia and the Pacific) and Uto-Aztecan (the native vernaculars of the Americas). These four groups account for more than a third of the 7,000 or so tongues spoken around the world today.

For each family, Dr Dunn and his team identified sets of cognates. These are etymologically related words that pop up in different languages. One set, for example, contains words like "night", "Nacht" and "nuit". Another includes "milk" and "Milch", but not "lait". The result is a multidimensional Venn diagram that records the overlaps between languages.

Which is fine for the present, but not much use for the past. To substitute for fossils, and thus reconstruct the ancient branches of the tree as well as the modern-day leaves, Dr Dunn used mathematically informed guesswork. The maths in question is called the Markov chain Monte Carlo (MCMC) method. As its name suggests, this spins the software equivalent of a roulette wheel to generate a random tree, then examines how snugly the branches of that tree

fit the modern foliage. It then spins the wheel again, to tweak the first tree ever so slightly, at random. If the new tree is a better fit for the leaves, it is taken as the starting point for the next spin. If not, the process takes a step back to the previous best fit. The wheel whirrs millions of times until such random tweaking has no discernible effect on the outcome.

When Dr Dunn fed the languages he had chosen into the MCMC casino, the result was several hundred equally probable family trees. Next, he threw eight grammatical features, all related to word order, into the mix, and ran the game again.

The results were unexpected. Not one correlation persisted across all language families, and only two were found in more than one family. It looks, then, as if the correlations between grammatical features noticed by previous researchers are actually fossilised coincidences passed down the generations as part of linguistic culture. Nurture, in other words, rather than nature. If Dr Dunn is correct, that leaves Dr Chomsky's ideas in tatters, and raises questions about the very existence of a language organ. You may be sure, though, that the Chomskyan heavy artillery will be making its first ranging shots in reply, even as you read this article. Watch this space for further developments.

 From the print edition: Science and technology