

A Turkish origin for Indo-European languages

Disease-mapping methods add geographical history to language family tree.

Alyssa Joyce

23 August 2012

Languages as diverse as English, Russian and Hindi can trace their roots back more than 8,000 years to Anatolia — now in modern-day Turkey. That's the conclusion of a study¹ that assessed 103 ancient and contemporary languages using a technique normally used to study the evolution and spread of disease. The researchers hope that their findings can settle a long-running debate about the origins of the Indo-European language group.

English, Dutch, Spanish, Russian, Greek and Hindi might all sound very different, but there are many commonalities, such as the Dutch *moeder*, Spanish *madre* and Russian *mat'*, all of which mean "mother". On this basis, researchers have concluded that more than a hundred languages across Europe and the Middle East, from Iceland to Sri Lanka, stem from a common ancestor.

Some scholars think that Indo-European languages spread with farming techniques from Turkey across Europe and Asia 8,000–9,500 years ago². Others suggest that nomadic 'Kurgan' horsemen brought the origins of Indo-European language from central Asia about 6,000 years ago³. There is archaeological evidence to support both theories, but genetic studies of Indo-Europeans have been inconclusive, leading to an intractable debate among linguists, anthropologists and cultural historians.

Taking sides

In 2003, Russell Gray and his then doctoral student, Quentin Atkinson, at the University of Auckland in New Zealand generated a maelstrom of controversy by claiming to have solved, by computer modelling, what has been described as "the most intensively studied, yet still the most recalcitrant, problem of historical linguistics", coming down on the side of Anatolia⁴ (see '[Language tree rooted in Turkey](#)').

Neither Gray nor Atkinson is a linguist. But they believed they could work with the kinds of tools employed in evolutionary ecology to answer important questions about language prehistory.

Genes and words have several similarities, and language evolution has conventionally been mapped using a "family tree" format. Gray and Atkinson theorized that the evolution of words was similar to the evolution of species, and that the 'cognate' of words — how closely their sounds and meanings are related to one another — could be modelled like DNA sequences and used to measure how languages evolved. By extension, the rate at which words changed — or mutated — could be used to determine the age at which Indo-European languages diverged from one another.

Using methods from evolutionary biology, the duo compared common words in 87 Indo-European languages, such as 'mother', 'hunt' and 'sky', to figure out how language 'species' were related to one another⁴. They traced the origins of Indo-European languages to 7,800–9,800 years ago, supporting the Anatolian hypothesis.

Critics were sceptical. Gray and Atkinson had determined when the languages originated, but not where. So, in a paper published today in *Science*, Atkinson, Gray and their colleagues address this using the type of geography-based computer modelling normally used by epidemiologists to track the spread of disease¹.



Jochen Tack/Getty

Indo-European languages appear to have originated around 7,800 to 9,800 years ago in Turkey.



The locations of current Indo-European languages are well known, and the geographic origin of older, extinct languages — such as Ancient Greek or Sanskrit — can be inferred from the historical record. In this way, the researchers believed they could track the movement of the Indo-European languages in the same way that epidemiological models trace a disease outbreak to its source. Once again, they conclude that the origin is Anatolia.

A clear spatial picture

“Finally we have a clear spatial picture,”

says Colin Renfrew at the University of Cambridge, UK, who originally proposed Anatolia as the source of the Indo-European language family. But he predicts that many historical linguists will be slow to accept the evidence. “The structure of ‘Indo-European studies’ has been founded for so long on the myth of mounted Kurgan warrior horsemen riding down from the Russian steppes that it will take scholars a while to recover,” he says.

Indeed, many linguists and archaeologists still favour the Kurgan hypothesis. Andrew Garrett, a linguist at the University of California, Berkeley, considers the new methods innovative, but he remains unconvinced. “There is bias in the underlying data that leads to an erroneous conclusion, and strong evidence that is ignored which still strongly supports the Kurgan hypothesis,” he says. David Anthony, an archaeologist at Hartwick College in Oneonta, New York, says this type of model doesn’t match the complex linguistic and archaeological evidence. “The study is an example of retrofitting evidence to a model, but the results of such a model are only as useful as the underlying data and assumptions,” he says.

But Atkinson says the new models are slowly becoming more accepted in the field. “Ten years ago, responses to this work were very different. I have noticed a real shift in attitudes towards computational-modelling approaches in historical linguistics, from being just an odd sideshow to a clear focus of attention.”

Nature | doi:10.1038/nature.2012.11270

References

1. Bouckaert, R. *et al. Science* **337**, 957–960 (2012).
2. Renfrew, C. *Archaeology and Language: The Puzzle of Indo-European Origins* (1990).
3. Gimbuta, M. *The Prehistory of Eastern Europe* (1956).
4. Gray, R. D. & Atkinson, Q. D. *Nature* **426**, 435–439 (2003).