New model set to reveal oldest spoken words yet

February 10 2015

A team of University of Reading scientists has developed a mathematical technique that can work out when changes to how words are pronounced occurred in different languages.

The model will give researchers an opportunity to discover the earliest words and languages spoken to date, with the potential to go back thousands of years

Led by Professor Mark Pagel, who worked with colleagues from the Santa Fe Institute in the USA, the technique detects historical 'concerted sound changes', a phenomenon in languages where a specific sound changes to the same other sound simultaneously in many words. For instance, the English father and foot replacing the p sound in the Latin roots of pater and ped with an f sound.

The model was tested on the evolution of Turkic, a language family of at least thirty-five languages, spoken by Turkic peoples from Southeastern Europe and the Mediterranean, to Siberia and Western China.

It identified more than 70 regular sound changes that occurred throughout the 2000 year history of the Turkic languages. These include the word 'pas' (meaning 'head') in the Khakassian language. The initial 'p' has changed to 'b' yielding 'baš' in Turkish, Uzbek and 16 other Turkic languages. Similarly, 'pel-' (meaning 'louse') in Khakassian becomes 'bil-' or 'bel' in the other languages.

Professor Pagel, a leading evolutionary biologist at the University of Reading said: "Our new method is another exciting step to understanding how languages and genes evolve. It will allow us to go back in time further than before, making it possible to reconstruct ancient proto-languages, words that might have been spoken many thousands of years ago.

Professor Pagel: "Intriguingly, this concerted linguistic change has a parallel in genetics where the same changes can happen to several different genes simultaneously. The model provides a tool that genetic researchers can use to identify concerted evolutionary change in genes, such as the evolution of the mammalian sex-chromosomes."

The model boosts Reading researchers' cutting-edge work which is lifting the lid on the origins of language and how languages evolved over time. Professor Pagel's previous research on the evolution of human languages has built up a fascinating picture of how our 7,000 living human languages have evolved.

The research team has documented the shared patterns in the way we use language and discovered why some words succeed and others have become obsolete over time. The team has also found that Ice Age people living in Europe 15,000 years ago might have used forms of some common [words](#) that in some cases could still be recognised today.

The team use statistical estimates of rates of lexical replacement for a range of vocabulary items in the Indo-European languages. The variation in replacement rates makes the most common vocabulary items in these languages promising candidates for estimating the divergence between pairs of languages.

The paper Detecting Regular Sound Changes in Linguistics as Events of Concerted Evolution, was published last month in the journal *Current*

Biology.

More information: "Detecting Regular Sound Changes in Linguistics as Events of Concerted Evolution" DOI: dx.doi.org/10.1016/j.cub.2014.10.064