Study suggests Sino-Tibetan language family originated in present-day northern China

25 April 2019, by Bob Yirka

A small team of researchers affiliated with several institutions across China has conducted an in-depth study of the Sino-Tibetan language family and has concluded that it likely originated in present-day northern China. In their paper published in the journal *Nature*, the group describes their research efforts and the conclusions they drew from them. Randy LaPolla with Nanyang Technological University, in Singapore, has published a News and Views piece describing the work by the team in the same journal issue.

Most scientists who study language and its history agree that most of the languages that are spoken today across the globe originated from much smaller subsets. The Romance, Germanic and Slavic languages, for example, are believed to have originated from a proto-language that is now labeled as the Indo-European family of languages—this family represents the most spoken language family in the world today. The second most spoken language family is called the Sino-Tibetan language family and it includes Mandarin, Cantonese, Tibetan and approximately 400 other languages. This second largest group represents approximately 20 percent of people alive today, or one and a half billion people. In this new effort, the researchers sought to settle a longstanding debate in the linguistics community—did the Sino-Tibetan language family originate in northern or southern China, or perhaps even somewhere else. And when did it come to exist?

To find the answer, the researchers applied a wide variety of tools to the problem, from multiple fields. Their approach involved building an evolutionary tree similar to that used in the biology fields—from the top down. The branches at the top would represent languages spoken today and the base would represent the Sino-Tibetan proto-language. To build their tree, the researchers used genetic and linguistic data and borrowed tools from computational biology and anthropology. They also used historical information such as groups of people migrating. During the later stages of their work, the team applied probability testing to help build the branches on their tree.

In the end, the researchers found that the Sino-Tibetan language family likely originated in northern parts of China and spread south and west as people moved to farm new lands. They also suggest the language family likely got its start prior to 6000 years ago—their tree showed that time frame as the point at which the first branches began to appear.
