

Origin of Sino-Tibetan language family revealed by new research

May 6 2019



Harvesting foxtail millet (*Setaria italica*) in Taiwan. Credit: Chih-hung Yang

The Sino-Tibetan language family includes early literary languages, such as Chinese, Tibetan and Burmese, and is represented by more than 400 modern languages spoken in China, India, Burma, and Nepal. It is one of

the most diverse language families in the world, spoken by 1.4 billion speakers. Although the language family has been studied since the beginning of the 19th century, scholars' knowledge of the origin of these languages is still severely limited. An interdisciplinary study published in *PNAS*, led by scientists of the Centre des Recherches Linguistiques sur l'Asie Orientale (Paris), the Max Planck Institute for the Science of Human History (Jena), and the Centre de Recherches en Mathématiques de la Décision (Paris), now sheds new light on the place and date of the origin of these languages. Based on a phylogenetic study of 50 ancient and modern Sino-Tibetan languages, the scholars conclude that the Sino-Tibetan languages originated among millet farmers, located in North China, around 7,200 years ago.

During the past 10,000 years, two of the world's largest language families emerged, one in the west and one in the east of Eurasia. Together, these families account for nearly 60 percent of the world's population: Indo-European (3.2 billion speakers), and Sino-Tibetan (1.4 billion). The Sino-Tibetan family comprises about 500 languages spoken across a wide geographic range, from the west coast of the Pacific to Nepal, India, and Pakistan. Speakers of these languages have played a major role in human prehistory, giving rise to early high cultures China, Tibet, Burma and Nepal. However, while archaeogeneticists, phylogeneticists, and linguists have energetically discussed the origins of the Indo-European language family, the formation of Sino-Tibetan languages has previously received little attention.

One of the world's most diverse language families

"The Sino-Tibetan language family is one of the most diverse families in the world. It includes all of the different types of morphological systems, ranging from isolating languages, such as Chinese, Burmese, and Tujia, to polysynthetic languages, such as Gyalrongic and Kiranti languages," explains Guillaume Jacques of the Centre des Recherches Linguistiques

sur l'Asie Orientale, co-first author of the study. "While our knowledge of how to compare these languages linguistically is improving, important aspects of the development of their sound systems and their grammar remain poorly understood."

A database of core words from 50 Sino-Tibetan languages

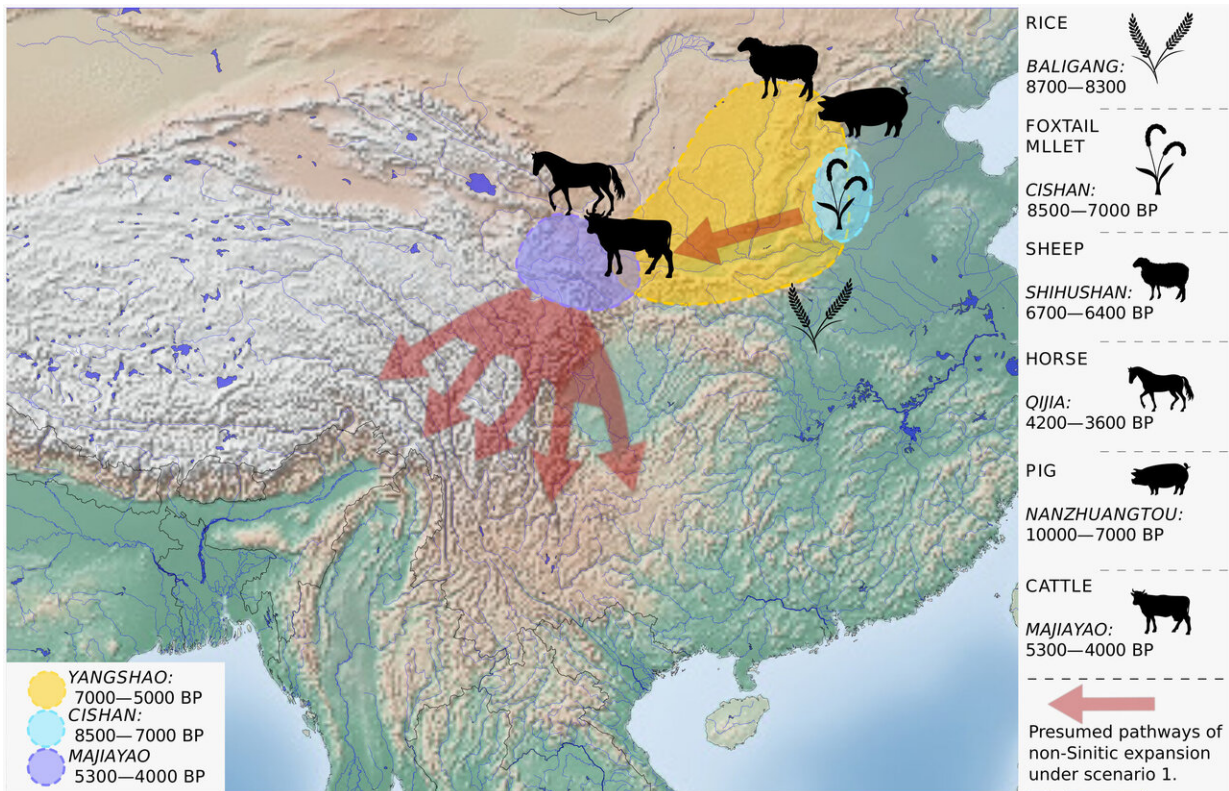
In order to shed light on the complex history of these languages, the scholars assembled a lexical database containing core vocabulary from 50 Sino-Tibetan languages. This database, published here for the first time, includes ancient languages spoken 1000 and more years ago, such as Old Chinese, Old Burmese, and Old Tibetan, as well as [modern languages](#) documented by field work.

"In order to compare these languages in a transparent way, we developed a specific annotation framework that allows us not only to mark which words we identify as sharing a common origin, but also which sounds in the words we think are related," says Johann-Mattis List of the Max Planck Institute for the Science of Human History, who led the study. "A particular problem in identifying the truly related words were the numerous cases where languages borrowed words from each other," mentions Jacques. "Luckily, we know the history of particular languages rather well and could rely on techniques that we developed before to reveal the true history concealed by these borrowings."

Evolutionary trees suggest that the language family originated about 7200 years ago

Using powerful computational phylogenetic methods, the team inferred the most probable relationships between these languages and then estimated when these languages might have originated in the past. "We find clear evidence for seven major subgroups with a complex pattern of

overlapping signals beyond that level," says Simon J. Greenhill of the Max Planck Institute for the Science of Human History. "Our estimates suggest that the ancestral [language](#) has arisen around 7,200 years ago."



Presumed pathways of the expansion of non-Sinitic Sino-Tibetan languages, contrasted with findings of early domesticates and early Neolithic cultures in China. Credit: J.-M. List and H. Sell

An agricultural analysis reveals the most likely origin and expansion scenario of the language family

To further resolve the complex pathways of the evolution of the Sino-Tibetan languages, the authors looked at related words describing

domesticates, because they may reveal how agricultural knowledge spread through the region. This agricultural analysis suggests an origin of the Sino-Tibetan [family](#) in Northern Chinese communities of millet farmers of the Neolithic cultures of late Cishan and early Yangshao.

"The most likely expansion scenario of the languages involves an initial separation between an Eastern group, from which the Chinese dialects evolved, and a Western group, which is ancestral to the rest of the Sino-Tibetan languages," summarizes Laurent Sagart of the Centre des Recherches Linguistiques sur l'Asie Orientale, co-first author of the study, who carried out the agricultural analysis.

"We are very excited about our findings," says List. "Our approach combines robust, traditional scholarship with cutting-edge computational methods within a computer-assisted framework that allows us to use our knowledge of today's languages as a key to their past."

More information: Laurent Sagart et al., "Dated language phylogenies shed light on the ancestry of Sino-Tibetan," *PNAS* (2019).

www.pnas.org/cgi/doi/10.1073/pnas.1817972116

Provided by Max Planck Society

Citation: Origin of Sino-Tibetan language family revealed by new research (2019, May 6) retrieved 27 April 2024 from

<https://phys.org/news/2019-05-sino-tibetan-language-family-revealed.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--