Do all languages come from the same source? Author: Allan Bomhard

Are all languages related? Do they all come from the same source?

Have you ever studied German? Or Spanish, or French? If you have, you were probably grateful for cognates, foreign words that sound and look like English words with related meanings. In German, your parents are your Mutter and your Vater. In Spanish, they're your madre and padre. In French, they're your mère and père.

These resemblances not only make language learning easier, they tell us something about the history of languages. English and German share some similar vocabulary because they're both descendants of a language called Proto-West-Germanic, spoken by tribes in northern Europe well over two thousand years ago. Over time, migrations split that language into dialects, and some of the tribes moved across the North Sea into the British Isles. Fifteen centuries of separate development turned the speech of the islands into varieties of English. And the language of the Mainlanders turned into varieties of German. So we have two languages, obviously different, but also so alike that they're clearly part of the same family.

Language families, like families of people, can be connected into larger and larger groupings, spreading outward and backward in time, as your relatives do on a genealogy chart. The Germanic family that English, German and several other languages belong to has a cousin, the Romance family, which includes not only French and Spanish but also several other languages that have Latin as their common ancestor.

Now let's go back a step further in time. The Germanic and Romance families share a common ancestor called Proto-Indo-European. It was spoken by tribes living some six to seven thousand years ago, probably in the steppes north and east of the Black Sea. From there the tribes spread westward across Europe and eastward and southward into Iran and northern India. As they spread and lost contact with each other, their language changed into languages like Greek, Armenian, and Albanian, and into families like Germanic, Romance, Celtic, Slavic, Indian languages, and Iranian languages. Taken together, they make up the Indo-European family, the most widely-spoken languages in the world today.

Different as the Indo-European languages were from one another, they all preserved bits of ancient vocabulary and grammar. And linguists have used these bits to figure out relationships and actually reconstruct the older languages. Sir William Jones opened the way in the 18th century through a remarkable analysis of the classical Indian language Sanskrit, showing that Sanskrit was related to languages in Europe. And now, even though no one has seen or spoken the original Indo-European language for thousands of years, we know what it looks like. And by cracking the code of Indo-European, we've taken a big step toward answering the question: can all language families be linked in a super family tree that begins with a single ancestral language?

To find out, linguists have increasingly studied and compared non-Indo-European languages, asking: What families do they belong to? How far back can those families be traced? Clearly, many non-Indo-European languages can be grouped together. I've always found it amazing that Finnish, Estonian, and Hungarian -- surrounded by Indo-Europeans in the heartland of Europe, are not in that family. But they do group together with several other languages to form a non-Indo-European family called Uralic. And there's a family called Turkic, which takes in Turkish, Azerbaijan, Uzbek, Uighur, and some other languages in central Asia. And in China, the Sino-Tibetan family includes over 250 languages. We think that at least 200 families exist, but are they related?

There are theorists who believe we can lump all the languages of the world -- including even oddballs like Basque, that seem to fit nowhere -- into a handful of giant families.

But maybe we can't go that far. The fact that the word for "dog" in an Australian native language called Mbabaram is "dog" doesn't mean that Mbabaram is related to English; it's just a random resemblance. The fact that Chinese calls coffee "kafei" doesn't mean that Chinese is related to English, either; the origin is a Turkish word that happens to have been borrowed by both Chinese and English. And we know that languages change continuously; new words join the vocabulary and older words, including cognates, disappear. After tens of thousands of years of change, can we reliably find a common ancestor? Do all languages come from the same source? The answer is: Maybe... and maybe not. It's too soon to know.