

Chomsky was wrong: evolutionary analysis shows languages obey few rules

By [John Timmer](#) | Last updated about 2 hours ago

Human languages are far more complex than any animal communication system we're aware of, and yet young children can easily learn to master more than one language in an astonishingly short period of time. This has led a number of linguists, most notably Noam Chomsky, to suggest that there might be language universals, common features of all languages that the human brain is attuned to, making learning easier. Now, a team of cognitive scientists has teamed up with an evolutionary biologist to perform a phylogenetic analysis of language families, and the results suggest that when it comes to languages, there are no rules.

The authors of the new paper point out just how hard it is to study languages. We're aware of over 7,000 of them, and they vary significantly in complexity. There are a number of large language families that are likely derived from a single root, but a large number of languages don't slot easily into one of the major groups. Against that backdrop, even a set of simple structural decisions—does the noun or verb come first? where does the preposition go?—become dizzyingly complex, with different patterns apparent even within a single language tree.

Linguists, however, have been attempting to find order within the chaos. Noam Chomsky helped establish the Generative school of thought, which suggests that there must be some constraints to this madness, some rules that help make a language easier for children to pick up, and hence more likely to persist. Others have approached this issue via a statistical approach (the authors credit those inspired by Joseph Greenberg for this), looking for word-order rules that consistently correlate across language families. This approach has identified a handful of what may be language universals, but our uncertainty about language relationships can make it challenging to know when some of these are correlations are simply derived from a common inheritance.

For anyone with a biology background, having traits shared through common inheritance should ring a bell. Evolutionary biologists have long been able to build family trees of related species, called phylogenetic trees. By figuring out what species have the most traits in common and grouping them together, it's possible to identify when certain features have evolved in the past. In recent years, the increase in computing power and DNA sequences to align has led to some very sophisticated phylogenetic software, which can analyze every possible tree and perform a Bayesian statistical analysis to figure out which trees are most likely to represent reality.

By treating language features like subject-verb order as a trait, the authors were able to perform this sort of analysis on four different language families: 79 Indo-European languages, 130 Austronesian languages, 66 Bantu languages, and 26 Uto-Aztecan languages. Although we don't have a complete roster of the languages in those families, they include over 2,400 languages that have been evolving for a minimum of 4,000 years.

The results are bad news for universalists: “most observed functional dependencies between traits are lineage-specific rather than universal tendencies,” according to the authors. The authors were able to identify 19 strong correlations between word order traits, but none of these appeared in all four families; only one of them appeared in more than two. Fifteen of them only occur in a single family. Specific predictions based on the Greenberg approach to linguistics also failed to hold up under the phylogenetic analysis. “Systematic linkages of traits are likely to be the rare exception rather than the rule,” the authors conclude.

If universal features can't account for what we observe, what can? Common descent. “Cultural evolution is the primary factor that determines linguistic structure, with the current state of a linguistic system shaping and constraining future states.”

It's important to emphasize that this study looked at a specific language feature (word order). Although that's a fairly significant one, it still leaves a lot of areas open for linguists to argue about. And the study did not build an exhaustive tree of any of the language families, in part because we probably don't have enough information to classify all of them at this point.

Still, it's hard to imagine any further details could overturn the gist of things, given how badly features failed to correlate across language families. And the work might be well received in some communities, since it provides an invitation to ask a fascinating question: given that there aren't obvious word order patterns across languages, how does the human brain do so well at learning the rules that are a peculiarity to any one of them?

Nature, 2011. DOI: [10.1038/nature09923](https://doi.org/10.1038/nature09923) ([About DOIs](#)).