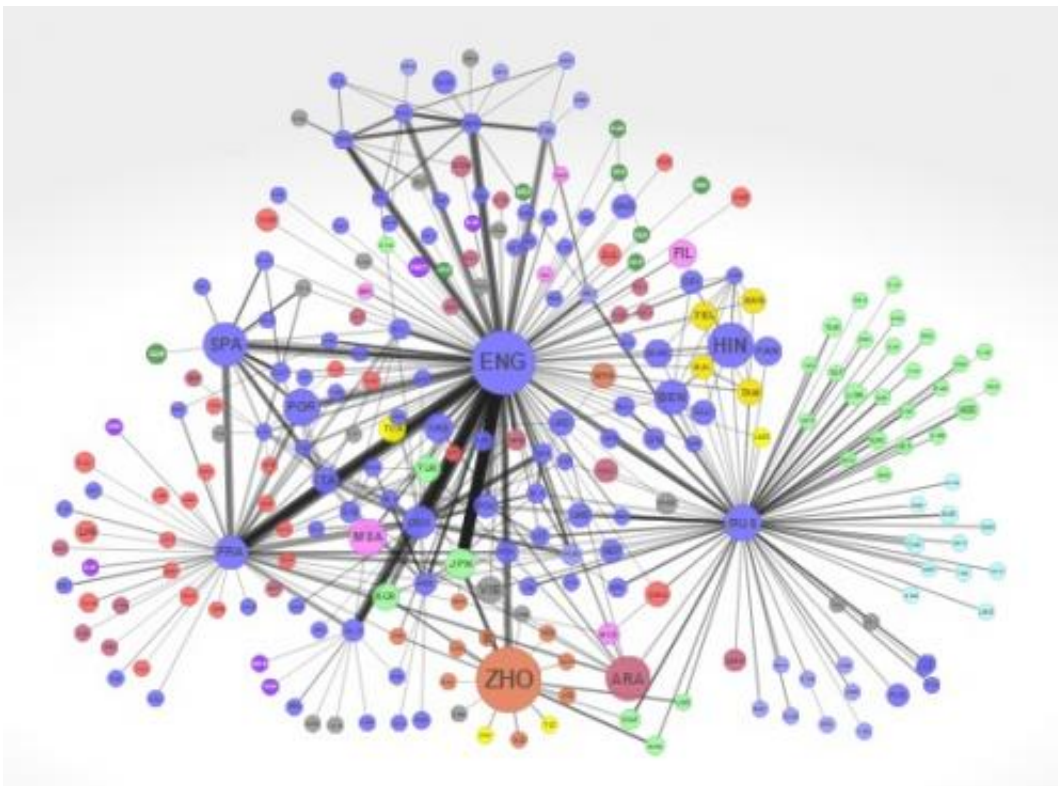


Your chance of becoming globally famous depends on the language you speak

December 16 2014, by Lisa Zyga



The global language network for book translations. Node sizes are proportional to the number of speakers (native plus nonnative) of each language. Node colors indicate language families, and link colors show the significance of the link. Link widths show the total number of translations. Credit: Ronen, et al.

(Phys.org)—Is there a connection between language and fame? A recent study has found that the number of famous people a country produces is

more strongly correlated to that country's language than to its wealth or population. So a person born in an English-speaking country, where the language has a large global influence, has a greater chance of becoming famous than someone born in a country in which the language is less globally influential.

This correlation between language and fame is just one result gleaned from the creation of a new global language network. In the new study published in *PNAS*, researchers led by César A. Hidalgo at MIT have mapped out global language networks in order to measure a language's centrality, from which they can extract new insights in a variety of areas.

To do this, the researchers compiled millions of pieces of data in which a piece of written text was translated from one language to another—a feat that has become possible only in the past few years due to large online data records and the software to analyze it. The researchers used three data sources: 2.2 million book translations from UNESCO's *Index Translationum* project; 382 million Wikipedia edits, where users often made edits to more than one Wikipedia language edition; and 550 million tweets from users who tweeted in more than one language. See the interactive networks [here](#).

To measure the centrality of a language in each of these networks, the researchers used a tool called eigenvector centrality, which is also the basis for Google's PageRank algorithm. This method accounts for not only the connectivity of the language in question, but also that of its neighbors and its neighbors' neighbors, in an iterative manner.

The three global language networks derived from these three data sets are strongly correlated in several ways. All three networks show English as the most central hub, along with a handful of intermediate hub languages, including Spanish, German, and French. Some languages, such as Chinese, Arabic, and Hindi, may be spoken by very large

numbers of people, yet are more peripheral in the network due to the low volume of translations between them and the hub languages. This finding supports the well-known problem that the low number of translations into Arabic is a major obstacle in disseminating outside knowledge into the Arab world.

In other ways, the three networks are somewhat different. For instance, the Twitter and Wikipedia datasets exhibit a larger share of languages associated with developing countries, such as Malay, Filipino, and Swahili, compared to the written books dataset. This result suggests that the newer, less formal channels of communication are more inclusive of populations in developing countries, compared to written books.

The eigenvector centrality method also formalizes the intuitive idea that more influential languages provide more direct paths of translations to other languages. For example, the researchers explain that it is easy for an idea conceived by a Spanish speaker to directly reach an English speaker through bilingual speakers of English and Spanish. However, it is more difficult for an idea conceived by a Vietnamese speaker to directly reach a Mapudungun speaker in Chile because far fewer people are bilingual in both Vietnamese and Mapudungun. Instead, the idea might travel from Vietnamese to English to Spanish to Mapudungun.

It also makes sense that better connected languages should increase the visibility of the content produced by the speakers of that language. With this in mind, the researchers wanted to see how closely the eigenvector centrality of a language is correlated to the number of [famous people](#) who were born into that language. Their list of famous people (born between 1800 and 1950) comes from two sources: pantheon.media.mit.edu (an MIT project that maps cultural production throughout history) and the book *Human Accomplishment*.

The strong correlation between language and fame may not be that

surprising, but it is still impossible to tell from the data alone which is the cause and which the effect: Are the ideas produced in a hub language truly more noteworthy than ideas produced in other languages, causing more of these ideas to be translated into other languages? Or does a person born into a hub language have a greater chance of becoming famous because hub languages promote better visibility of their ideas?

The researchers suggest that the two mechanisms are not mutually exclusive, as they are likely to reinforce each other over time. So a language with high centrality may signal an abundance of earlier achievements by its speakers, and this rich history has increased the centrality of that language, enhancing the visibility of ideas produced by its current speakers.

In the future, assessments of changes in the structure of the global language networks can reveal important trends, such as whether English is gaining or losing influence with respect to rising powers such as India and China, or whether certain languages are heading toward extinction. In this way, the global language networks complement current predictions of language changes, which rely mostly on the [language's](#) number of speakers.

More information: "Links that speak: The global language network and its association with global fame," by Shahar Ronen et al. *PNAS*, www.pnas.org/cgi/doi/10.1073/pnas.1410931111

Watch a video of the researchers explaining similar work [here](#).

© 2014 Phys.org

Citation: Your chance of becoming globally famous depends on the language you speak (2014, December 16) retrieved 24 April 2024 from <https://phys.org/news/2014-12-chance-globally->

[famous-language.html](#)

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.