

F-bombs notwithstanding, all languages skew toward happiness

February 9 2015



University of Vermont scientists Peter Dodds and Chris Danforth led a new Big Data study, published in *PNAS*, that confirms the 1969 Pollyanna Hypothesis that there is a universal human tendency to "look on and talk about the bright side of life." Credit: Joshua Brown, UVM

In 1969, two psychologists at the University of Illinois proposed what they called the Pollyanna Hypothesis—the idea that there is a universal human tendency to use positive words more frequently than negative ones. "Put even more simply," they wrote, "humans tend to look on (and

talk about) the bright side of life." It was a speculation that has provoked debate ever since.

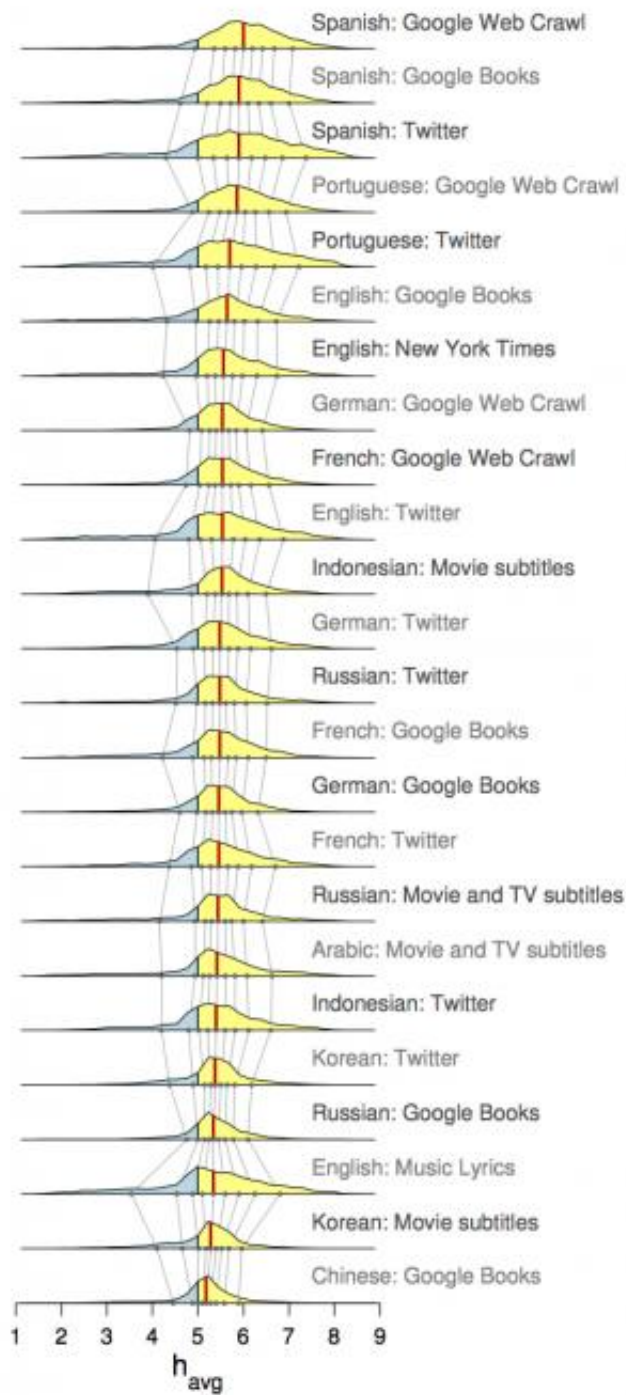
Now a team of scientists at the University of Vermont and The MITRE Corporation have applied a Big Data approach—using a massive data set of many billions of words, based on actual usage, rather than "expert" opinion—to confirm the 1960s guess.

Movie subtitles in Arabic, Twitter feeds in Korean, the famously dark literature of Russia, websites in Chinese, music lyrics in English, and even the war-torn pages of the *New York Times*—the researchers found that these, and probably all human language, skews toward the use of happy words.

"We looked at ten languages," says UVM mathematician Peter Dodds who co-led the study, "and in every source we looked at, people use more positive words than negative ones."

But doesn't our global torrent of cursing on Twitter, horror movies, and endless media stories on the disaster du jour mean this can't be true? No. This huge study of the "atoms of language—individual words," Dodds says, indicates that language itself—perhaps humanity's greatest technology—has a positive outlook. And, therefore, "it seems that positive social interaction," Dodds says, is built into its fundamental structure.

The new study, "Human Language Reveals a Universal Positivity Bias," appeared in the February 9 online edition of the *Proceedings of the National Academy of Sciences*.



Whether it's Arabic movie subtitles, books in Chinese, Spanish Twitter, German websites, or music lyrics in English -- a study published in the *Proceedings of the National Academy of Sciences*, by scientists at the University of Vermont and The Mitre Corporation, found a clear positive bias in human language. In other words, we -- humanity -- "use more happy words than sad words. Credit: Courtesy Dodds et al., *PNAS*

To deeply explore this Pollyanna possibility, the team of scientists at UVM's Computational Story Lab—with support from the National Science Foundation and The Mitre Corporation—gathered billions of words from around the world using twenty-four types of sources including books, news outlets, social media, websites, television and movie subtitles, and music lyrics. For example, "we collected roughly one hundred billion words written in tweets," says UVM mathematician Chris Danforth who co-led the new research.

From these sources, the team then identified about ten thousand of the most frequently used words in each of ten languages including English, Spanish, French, German, Brazilian Portuguese, Korean, Chinese (simplified), Russian, Indonesian and Arabic. Next, they paid native speakers to rate all these frequently-used words on a nine-point scale from a deeply frowning face to a broadly smiling one. From these native speakers, they gathered five million individual human scores of the words. Averaging these, in English for example, "laughter" rated 8.50, "food" 7.44, "truck" 5.48, "the" 4.98, "greed" 3.06 and "terrorist" 1.30.

A Google web crawl of Spanish-language sites had the highest average word [happiness](#), and a search of Chinese books had the lowest, but—and here's the point—all twenty-four sources of words that they analyzed skewed above the neutral score of five on their one-to-nine scale—regardless of the language. In every language, neutral words like "the" scored just where you would expect: in the middle, near five. And when the team translated words between languages and then back again they found that "the estimated emotional content of words is consistent between languages."

In all cases, the scientists found "a usage-invariant positivity bias," as they write in the study. In other words, by looking at the words people

actually use most often they found that, on average, we—humanity—
"use more happy words than sad words," Danforth says.

This new research study also describes a larger project that the team of fourteen scientists has developed to create "physical-like instruments" for both real-time and offline measurements of the happiness in large-scale texts—"basically, huge bags of words," Danforth explains.

They call this instrument a "hedonometer"—a happiness meter. It can now trace the global happiness signal from English-language Twitter posts on a near-real-time basis, and show differing happiness signals between days. For example, a big drop was noted on the day of the terrorist attack on Charlie Hebdo in Paris, but rebounded over the following three days. The hedonometer can also discern different happiness signals in US states and cities: Vermont currently has the happiest signal, while Louisiana has the saddest. And the latest data puts Boulder, CO, in the number one spot for happiness, while Racine, WI, is at the bottom.

But, as the new paper describes, the team is working to apply the hedonometer to explore happiness signals in many other languages and from many sources beyond Twitter. For example, the team has applied their technique to over ten thousand books, inspired by Kurt Vonnegut's "shapes of stories" idea. Visualizations of the emotional ups and downs of these books can be seen on the hedonometer website; they rise and a fall like a stock-market ticker. The new study shows that Moby Dick's 170,914 words has four or five major valleys that correspond to low points in the story and the hedonometer signal drops off dramatically at the end, revealing this classic novel's darkly enigmatic conclusion. In contrast, Dumas's Count of Monte Cristo—100,081 [words](#) in French—ends on a jubilant note, shown by a strong upward spike on the meter.

The new research "in no way asserts that all natural texts will skew

positive," the researchers write, as these various books reveal. But at a more elemental level, the study brings evidence from Big Data to a long-standing debate about human evolution: our social nature appears to be encoded in the building blocks of language.

More information: "Human language reveals a universal positivity bias," by Peter Sheridan Dodds et al. *PNAS*,
www.pnas.org/cgi/doi/10.1073/pnas.1411678112

Provided by University of Vermont

Citation: F-bombs notwithstanding, all languages skew toward happiness (2015, February 9)
retrieved 26 April 2024 from
<https://medicalxpress.com/news/2015-02-f-bombs-languages-skew-happiness.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>
--