

People Sure Tweet "Fuck" a Lot, Finds Science

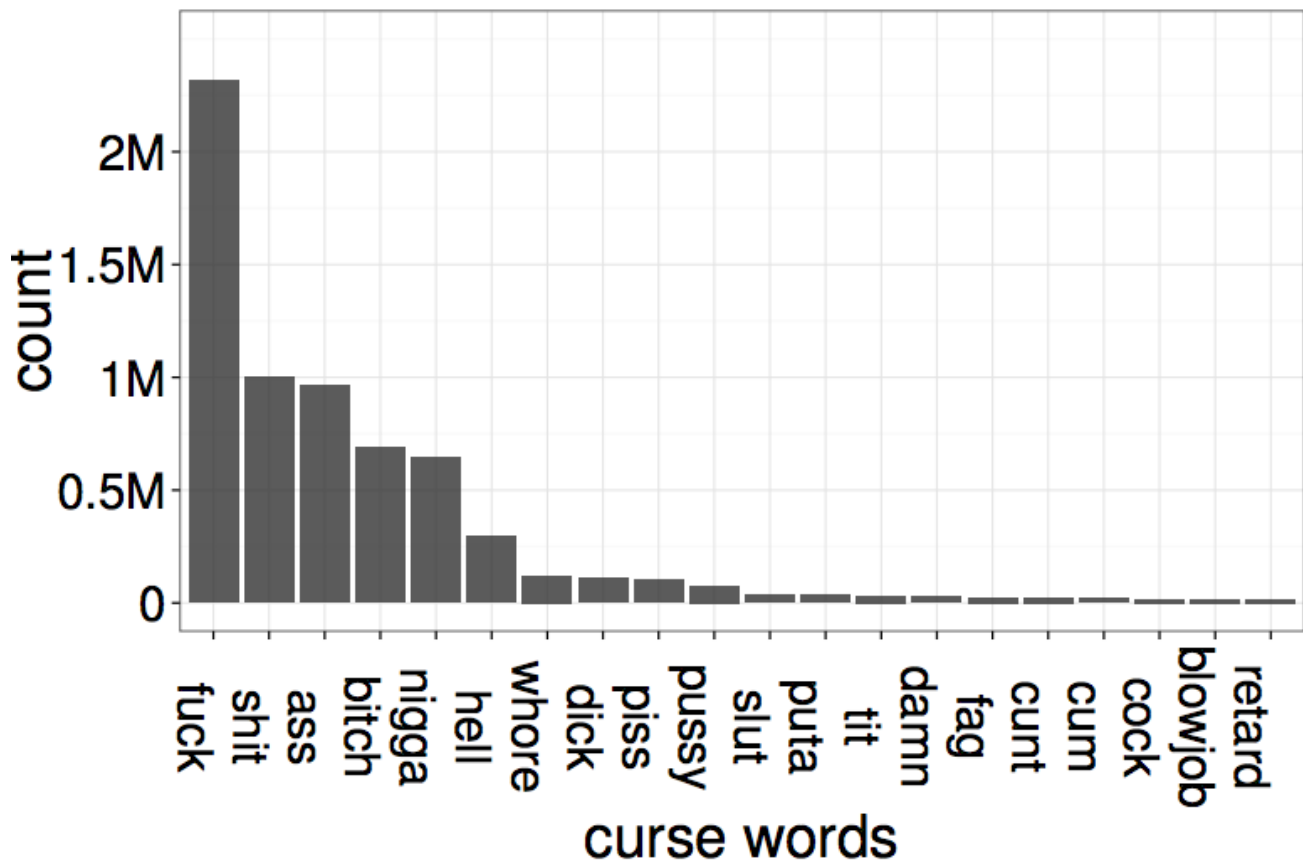
motherboard.vice.com/en_us/article/people-sure-tweet-fuck-a-lot-says-science

There are few things more satisfying than belting out a good “shit!” or “tarnation!” or whatever your swear word of choice is, but one of them surely is firing off a good, drunken rant into the ethereal therapist that is Twitter. Oh, this isn’t a confession, it’s science: According to a new language survey of Twitter, people swear a *lot* in their tweets, and the ratio of cursing to non-cursing tweets peaks between midnight and two AM.

The study analyzed about 51 million tweets from 14 million users to see if English speakers swear more often online than off, and more importantly, how those swear words are used. Frequency-wise, swearing is more prevalent on Twitter than in real life; the authors cite previous studies finding that “0.5% to 0.7% of all the words we speak in our daily lives are curse words,” while on Twitter the rate was 1.15 percent, or about double. (That’s after removing punctuation marks and emoticons from their data, the authors note. No word on middle finger ASCII art.)

According to Fast Co.Exist, the team of Wenbo Wang, Lu Chen, Krishnaprasad Thirunarayan, and Amit Sheth, all of whom hail from the Kno.e.sis Center at Wright State University, shared the results of the study this week at the ACM Conference on Computer-Supported Cooperative Work & Social Computing. The study is part of ongoing research into online language and harassment.

So if we swear more online—the authors also cite a 2006 study that found swear words made up 3 percent of words in online chat rooms—what choice curses do we use? Thankfully, there’s a graph for that:

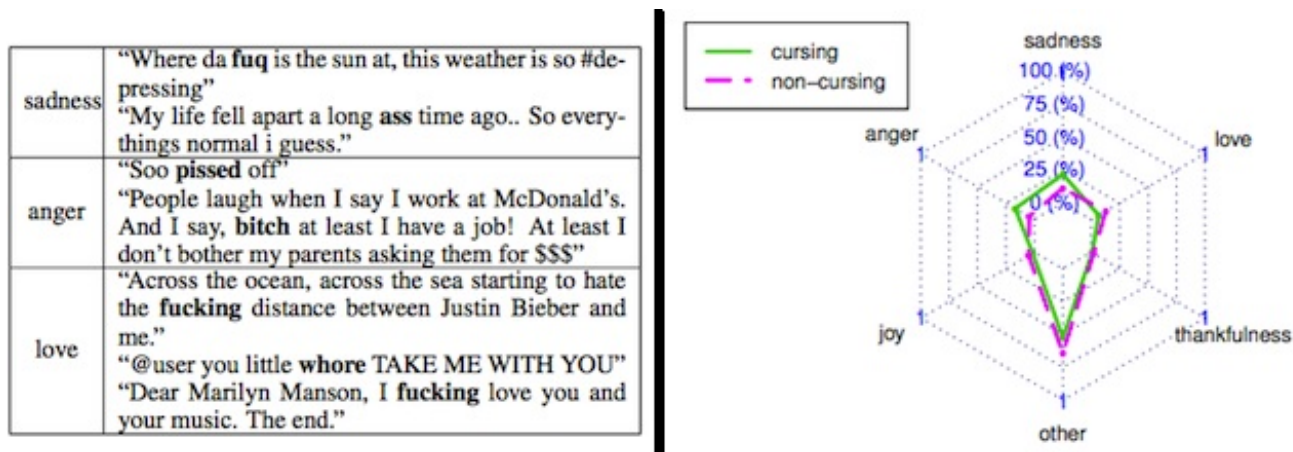


The frequency of the top 20 curse words in the study.

While it should come as no surprise that “fuck” is the star of the show, I didn’t expect that it would represent more than a third of all swear words said. And the swearing is really concentrated. According to the paper, “the top seven curse words – fuck, shit, ass, bitch, nigga, hell and whore cover 90.55% of all the curse word occurrences.” What, no love for damn or crap?

On its face, the finding that people swear more often on Twitter, where they’re less concerned about offending someone in public and where punchiness is key, is hardly a shock. Just as Johnny Carson said, at least according to [the Goodreads quotes site](#), “Never use a big word when a little filthy one will do.”

But looking at the actual words people use is telling. The team attempted to classify tweets according to emotion, which has been a popular area of [Twitter language research](#). While it’s not an easy task, the researchers found that “21.83% and 16.79% of the cursing tweets express sadness and anger emotions, respectively; in contrast, 11.31% and 4.50% of the non-cursing tweets express sadness and anger emotions, respectively.”

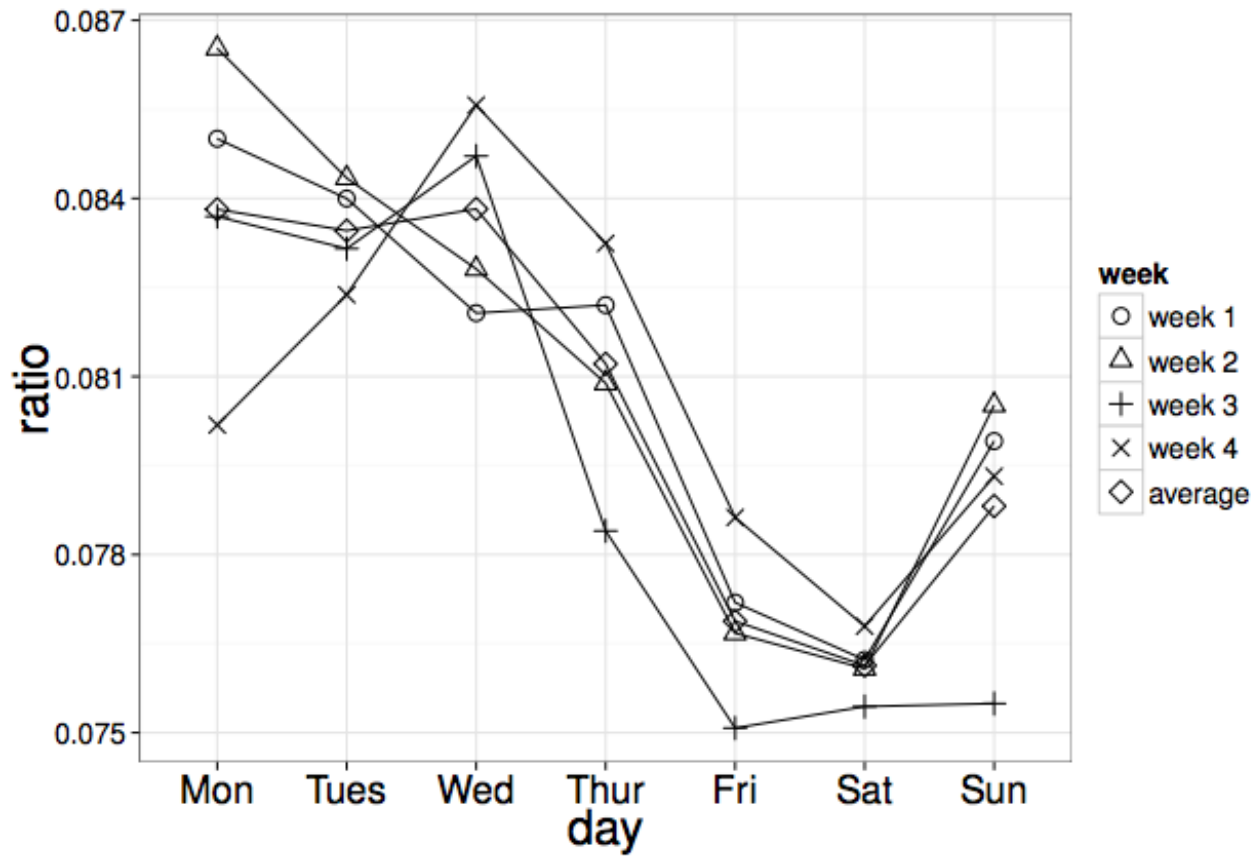


The table on the left shows how context can be used to ascribe emotion, but it's not easy. The graphic on the right show that while cursing tweets were more likely to be angry or sad, assigning emotions to many tweets remains difficult.

Looking at the most popular swear words, you can see how tough it is to figure out how they’re used in context. I’m sure all of the top seven words have been used in non-negative fashion by somebody, but three of them—“bitch,” “nigga,” and “whore”—stand out for their potential for harassment.

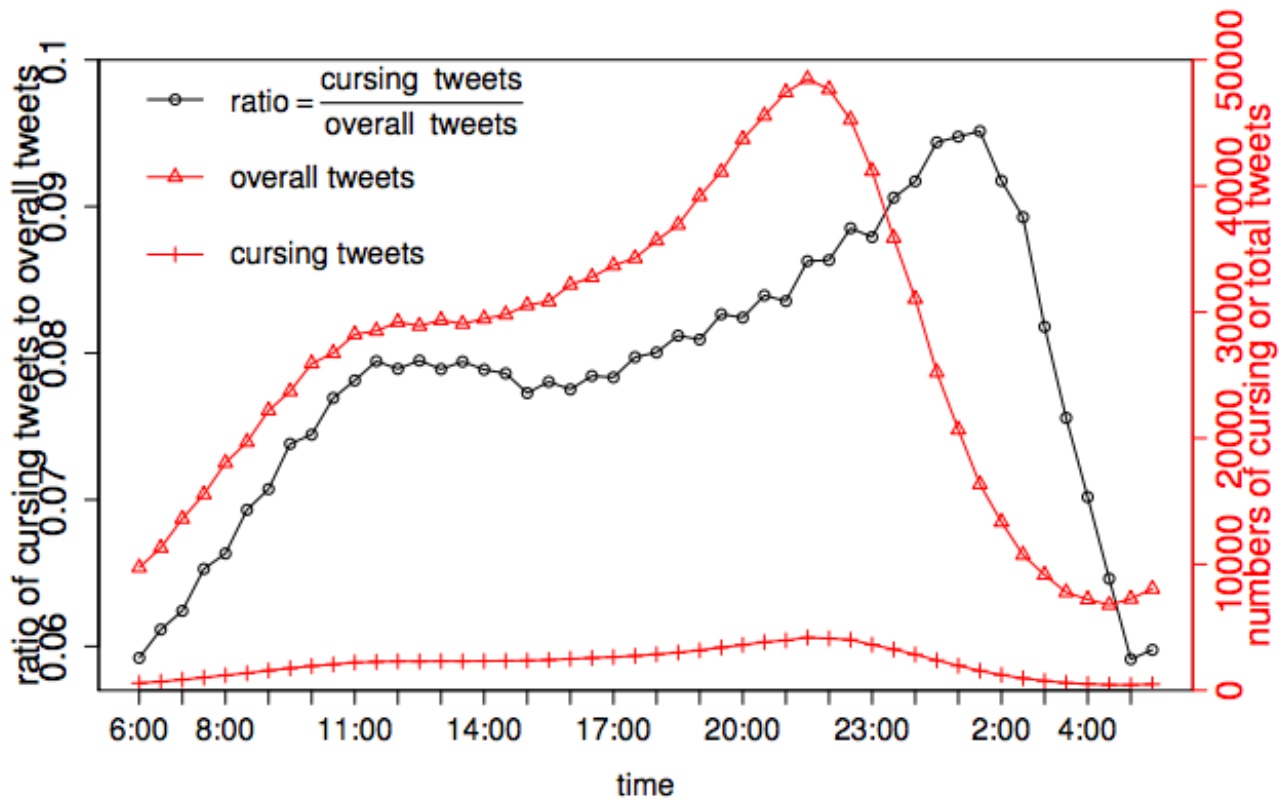
(It’s interesting that the researchers tried to count for multiple permutations of swear words—fuq, a\$\$, etc.—but don’t mention a distinction between “nigga” and “nigger,” which can be used in widely varying contexts. In fact, the researchers don’t discuss race much at all, which is surprising, especially considering the [sheer popularity of Twitter among African-Americans](#).)

So if we’ve established that people swear a lot on Twitter, when do they do it? Swearing volume is much higher in the first three days of the week, perhaps because Monday through Wednesday totally suck:



Swearing by day of the week.

Curiously, the ratio of swearing to non-swearing tweets peaks after midnight. Now, the researchers don't mention anything about people sucking back six gin and tonics before taking to the socialsphere, but from anecdotal evidence, I'd say late night Twitter does tend to get a bit more loose:



Cursing volume at different times of day.

So what does it all mean? Well, as the researchers note, ascribing motivation to such a large data set through an algorithm isn't completely precise. But looking at broader trends does produce some interesting conclusions. For one, considering how [much of a problem online harassment is](#), it's interesting to look at gender trends. According to the study, men swear the most, and swear most commonly at other men:

Sender	Recipient	Cursing Tweets (#)	Cursing Ratio (%)
F	M	3,808	3.81
F	F	3,977	3.98
M	F	4,192	4.19
M	M	5,483	5.48

Swearing trends by gender.

While finding solid gender demographic info for Twitter is difficult, [a 2013 Pew survey](#) found that male internet users are slightly more likely to use Twitter than females, which could explain the discrepancy. Still, that's fuzzy math, and the authors argue that the differences are pretty clear.

Specifically, men curse more than women, men overuse some curse words different from what women use and vice versa, and both men and women are more likely to curse in the same-gender contexts, the authors write.

At the end of the day, the study is more evidence that Twitter's open nature and vast repository of tweets is a great tool for language research, even if a whole lot of different groups are still working on [developing the perfect algorithm for analysis](#). But with researchers suggesting that social media analysis is powerful enough to [predict the](#)

next uprising or track the evolution of human language, it's work that has huge potential.

M