**[Science](https://www.theatlantic.com/science/)**

**The Randomness of Language Evolution**

English is shaped by more than natural selection.

[**Ed Yong**](https://www.theatlantic.com/author/ed-yong/)

Nov 1, 2017

Joshua Plotkin’s dive into the evolution of language began with clarity—and also a lack of it.

Today, if you wanted to talk about something that’s clear, you’d say that it has *clarity*. But if you were around in 1890, you would almost certainly have talked about its *clearness*.

[Plotkin](http://www.bio.upenn.edu/people/joshua-plotkin) first noticed this linguistic change while playing with Google’s Ngram Viewer, a search engine that charts the frequencies of words across millions of books. The viewer shows that a century ago, *clearness* dominated *clarity*. Now the opposite is true, which is strange because *clarity* isn’t even a regular form. If you wanted to create a noun from *clear*, *clearness* would be a more obvious choice. “Why would there be this big upswing in clarity?,” Plotkin wondered. “Is there a force promoting clarity in writing?”

It wasn’t clear. But as an evolutionary biologist, Plotkin knew how to find out.

The histories of linguistics and evolutionary biology have been braided together for as long as the latter has existed. Many of the earliest defenders of Darwinism were linguists who saw similarities between the evolution of languages and of species. Darwin himself wrote about these “curious parallels” in *The Descent of Man*. New words and grammatical rules are continually cropping up, fighting for existence against established forms, and sometimes driving those old forms extinct. “The survival ... of certain favored words in the struggle for existence is natural selection,” Darwin wrote.

Darwin, Plotkin says, used the way language changes “to popularize his heretical theory and explain for a broad audience what natural selection means. The process wasn’t easy to observe in organisms, but it was easier to see in words.”

But natural selection is just one force of evolutionary change. Under its influence, genes become more (or less) common because their owners are more (or less) likely to survive and reproduce. But genes can also change in frequency for completely random reasons that have nothing to do with their owner’s health or strength—and everything to do with pure, dumb luck. That process is known as drift, and it took decades for evolutionary biologists to recognize that it’s just as important for evolution as natural selection.

Linguists are still behind. It’s easy to see how languages can change through drift, as people randomly pick up the words and constructions they overhear. But when Darwin wrote about evolving tongues, he said, “The better, the shorter, the easier forms are constantly gaining the upper hand, and they owe their success to their own inherent virtue.” That’s a view based purely on natural selection, and it persists. “For the most part, linguists today have a strict Darwinian outlook,” Plotkin says. “When they see a change, they think there must be a directional force behind it. But I propose that language change, maybe lots of it, is driven by random chance—by drift.”

To see whether that was true, he and his colleagues [developed statistical tests](https://www.nature.com/nature/journal/vaop/ncurrent/full/nature24455.html) that could distinguish between the influence of drift and of natural selection. They then applied these tests to several online repositories, such the Corpus of Historical American English—a digital collection of 400 million words, pulled out of 100,000 texts published over the past 200 years.

The team focused first on the past-tense forms of verbs, and found at least six cases where natural selection is clearly in effect. In some cases, the verbs were regularized, losing weird past forms in favor of more-predictable ones that end in –*ed*. *Wove*, for example, gave way to *weaved*, while *smelt* lost ground to *smelled*. That’s not surprising: Many linguists have suggested that verbs tend to become more regular over time, perhaps because, like Darwin theorized, these forms are just easier to learn.

But Plotkin found just as many instances where selection drove verbs toward irregularity: *Dived* gave way to *dove*, *lighted* to *lit*, *waked* to *woke*, and *sneaked* to *snuck*. Why? Perhaps because we like it when words sound alike, and we change our language to accommodate such rhymes. For example, *dove* began to replace *dived* at the same time that cars became popular, and *drive*/*drove* became common parts of English. Similarly, the move from *quitted* to *quit* coincided with the rise of *split*, which became much more widely used when it acquired a new meaning—to leave or depart. In both cases, changes in one irregular verb—*drive* or *split*—may have irregularized others. “We can’t definitively say that’s the reason, but it’s coincident,” Plotkin says.

“It gets you to think harder about the motivation for change,” says [Salikoko Mufwene](http://mufwene.uchicago.edu/), from the University of Chicago. “The general claim is that there has been an evolution toward regularization, and they’re showing that this hasn’t always been the case. Now we need to think harder about when irregular forms are favored over regular variants.”

That is, if anything is favored at all. The team found that the changes that have befallen the vast majority of our verbs are entirely consistent with drift. You don’t need to invoke natural selection to explain why we say *spilled* instead of *spilt*, *burned* instead of *burnt*, and *knit* instead of *knitted*.

In other cases, drift and natural selection work together to shape languages. For example, Plotkin’s team also looked at the rapid rise of *do* in the 16th century, when phrases like “*You say not*” quickly changed into “*You do not say*.” They concluded that at first, the word randomly drifted its way into questions, so that “*Say you*?” gradually became “*Do you say*?” Once it became common, natural selection started pushing it into new contexts like declarative sentences, perhaps because it was easier for people to use it consistently.

The team also analyzed a third and more obscure grammatical change called Jespersen’s Cycle. In Old English, spoken before the Norman Conquest, speakers would negate a verb by putting a *not* in front of it. In Middle English, spoken between the 11th and 15th centuries, the negatives would surround the verb as they do in modern French (“*Je ne dis pas*”). And in Early Modern English, spoken between the 15th and 17th centuries, the negative followed the verb—the Shakespearean “I say not.” Now, we’ve come full circle, back to “*I don’t say*.”

Jespersen’s Cycle exists in many unrelated languages. In French, for example, the formal “*Je ne dis pas*” is giving way to the colloquial “*Je dis pas*.”

Natural selection still explains Jespersen’s Cycle far better than drift does, according to Plotkin's analysis. Perhaps it’s due to emphasis, he says. If one form is common, speakers could emphasize their disagreement by adding or subtracting words (“I don’t say that *at all*,” versus *“*I don’t say that”). As the emphatic forms become more common, they lose their sting, and are themselves replaced.

These results are part of a wider trend where linguists are starting to use these massive online corpora to address long-standing puzzles in language change. “This is an excellent trend,” says [Jennifer Culbertson](https://jennifer-culbertson.github.io/), from the University of Edinburgh. “Linguists have uncovered many really fascinating cases of language change, but the explanations on offer sometimes read like just-so stories. Random processes are simply underappreciated, because we want to come up with interesting explanations.” But by considering drift, too, linguists could “focus our energies on providing interesting explanations where they are really warranted.”

What about the change from *clearness* to *clarity*, which set Plotkin onto this quest in the first place? He says that he’s found signs of natural selection’s hand, but that will have to wait for another publication. “There’s lots to be done,” he says. “This is just the beginning of an investigation, which need not stop at written texts. Spoken records are just as ready and ripe for scrutiny.”

*We want to hear what you think about this article.* [*Submit a letter*](https://www.theatlantic.com/contact/letters/) *to the editor or write to letters@theatlantic.com.*

[**Ed Yong**](https://www.theatlantic.com/author/ed-yong/) is a staff writer at *The Atlantic*, where he covers science.