TIME

Redefining the Modern Dictionary

Katy Steinmetz @katysteinmetz May 12, 2016

Peaches Monroee was a young woman of no particular fame when she posted a video about her eyebrows on social network Vine two years ago. But that changed in the seconds it took her to describe those eyebrows as being “on fleek.” Spurred by her unfamiliar slang (and charm), the post went viral. Soon IHOP was tweeting about “pancakes on fleek.” Kim Kardashian West was using an #EyebrowsOnFleek hashtag to caption a selfie. Public curiosity soared. By late 2014, an average of 27,000 people were typing “What does on fleek mean?” into their Google search bars every month. About a year later, Dictionary.com got around to adding a definition: “flawlessly styled, groomed … looking great.”

The yearlong lapse—-which is relatively fast for a traditional dictionary to add a word—-highlights a problem that academics and linguists are trying to solve: lexicographers, the experts who write gold-standard definitions, cannot keep up with demand in an era when words travel by fiber-optic cable and not just by mouth or mail. As Aston University linguist Jack Grieve says, “A new word is more likely to take off and spread now than in the past because it’s easier for it to move.” Yet writing definitions remains a methodical exercise in observation and distillation. “We have to wait for words to settle both in meaning and pronunciation,” says Peter Sokolowski, editor-at-large for Merriam-Webster. “And those things take time.”

So while dictionary makers may rightly wait years before giving a new word the full treatment, other experts are trying to give the reading public something reliable in the meantime by finding a compromise between authority and speed. “A really good definition is a beautiful thing,” says Erin McKean, a lexicographer who founded the “Internet-age” dictionary company Wordnik. “I also believe that something is always better than nothing.” Other innovators are bringing rich data online for the first time and even reimagining the very nature of what a dictionary can be.

You may have heard of Urban Dictionary. Much like Wikipedia, that crowdsourced website can be a good place to start asking a question about what a word means, but experts say it doesn’t produce exhaustive answers–nor always serious ones. Take the top search result for on fleek: “A work [sic] used by uncultured idiots.” Says Dictionary.com lexicographer Jane Solomon: “You have to sift through a lot to get to these core pieces of information.” Lexicographer Jonathon Green, who wrote the most comprehensive dictionary of English slang ever printed, is less diplomatic: “It is a joke.”

Green parses even the most obscene, NSFW slang with the seriousness of a biologist dissecting a frog, and he plans to make his $625 set of three hulking volumes far more useful by turning

them into a searchable database, with 130,000 words and phrases, this summer. Green's Dictionary of Slang covers ground that traditional dictionaries will not, and that will help fill some gaps: as any lexicographer will tell you, even a word that you're encountering for the first time is probably older than you think. (Take this excerpt from a 1917 letter: "I hear that a new order of Knighthood is on the tapis–O.M.G.") At Green's London flat, it takes him 0.4 seconds to pull up 403 thoroughly researched synonyms for a promiscuous woman. "I feel the potential of my database is unlimited," he says.

But while going online will provide a portal for people to suggest new terms that need defining, Green wants the definitions to remain lexicographer-grade. If done well, a dictionary definition "stands in as a witness to the whole culture," as Oxford English Dictionary editor Edmund Weiner puts it. And that's a pretty high bar.

Lowering the bar is a key part of McKean's plan for Bay Area–based Wordnik, which aims to be more responsive than traditional dictionaries but more authoritative than crowdsourced sites. "There's a lot of word-shaped objects out there, way too many to write definitions for all of them," she says. "So how do we solve this problem? Well, it turns out journalists define words all the time." McKean raised nearly $60,000 on Kickstarter to invest in machine learning, so her startup can get better at scraping quality definitions from sources like magazines, blogs and newspapers. Soon, when people search for a word on her site, even if they don't get a definition, they will still get helpful clues: example sentences from articles, tweets that contain the word, pictures from Flickr that have been tagged with that term. Look up the word cookprint, for example, and up pops an excerpt from a Consumer Reports blog: "And cut down on your cookprint–the energy you use to prepare the food you eat." McKean thinks such snippets are helpful enough for many people's purposes. "Most of the words you learned in your life, you never read a definition for. You just read them in a sentence or heard someone say them," she says.

Then there are people like Roberto Navigli, a computer scientist and associate professor at Sapienza University in Rome, who is going far beyond the "what's that word mean" use case. He has built a network that may well be the dictionary of the future, one organized using the meanings of words, not their spellings. His database is part encyclopedia and part translator, in which terms come with illustrations–and will soon come with videos and animation. It includes entities as well as words, so a search for apple produces results that contain a picture of fruit as well as the famous corporate logo. Navigli calls it BabelNet, after the biblical tower and the technology he believes can bridge the world's languages. "The idea is to put a lot of resources together, all the resources that people usually access separately," he says. As it stands, BabelNet has 14 million entries, with information in 271 languages, drawing on more than a dozen giant sources of data.

The first two that Navigli combined were Wikipedia and a network built at Princeton University called WordNet. While print dictionaries organize words in relatively useless alphabetical order, the academics at Princeton manually arranged words based on their meanings, grouping sets that go, for instance, from general to specific and from part to whole. So there is a set that

connects vehicle to car to convertible and a connection from car to its components: the engine, the wheel, the piston. Like WordNet, BabelNet has the potential to help with artificial intelligence. While it's easy for humans to understand which bank a person means from context (whether it's a pool shot or the place to find an ATM), such distinctions–a concept known as disambiguation–aren't so simple for computers. If the likes of Siri hear the words money and bank in close proximity, that can help her figure out what a person is talking about.

When a user types a term into the BabelNet search bar, responses come back in a list that starts with the one it thinks you meant, partly based on how many connections the term has. Type in "plane" and up comes an airplane, followed by a geometric plane and a metaphysical plane. A user can hear the word spoken out loud and translate it into languages ranging from Haitian Creole to Cherokee. One can also explore the network, traveling from the airplane entry to amphibious aircraft to the history of aviation to the reclusive aviator Howard Hughes. To increase BabelNet's reliability, Navigli has released online games and this fall plans to launch a social network through which users will validate material that comes from places like Wikipedia.

And he's going to keep adding more data in the meantime, plowing ahead with all the optimism of a Silicon Valley technocrat. "The more we integrate, the more confirmation we have that some translation or some definition is appropriate. It is a virtuous cycle," Navigli says. Search for the meaning of that phrase on BabelNet and you get seven definitions, one of which is this: "a situation in which the solution to one problem makes each future problem easier to solve."

What's the word

Jack Grieve, a linguist at the U.K.'s Aston University, says that thanks to the web, new words can travel faster and existing ones can quickly become new obsessions. Here are a few examples:

[This article consists of 7 illustrations. Please see hardcopy of magazine or PDF.]

Baeless

(bey'-les)

Single; without a mate

On fleek

(awn fleek)

On point; looking good

Tookah

(took’-uh)

Marijuana; Mary Jane; weed; pot

Gainz

(geyns)

Weight gained through tough exercise

Famo

(fah’-moh)

Family and friends

Senpai

(sen-pahy’)

An older mentor

Cosplay

(kos’-pley)

Costumed role-playing

This appears in the May 23, 2016 issue of TIME.