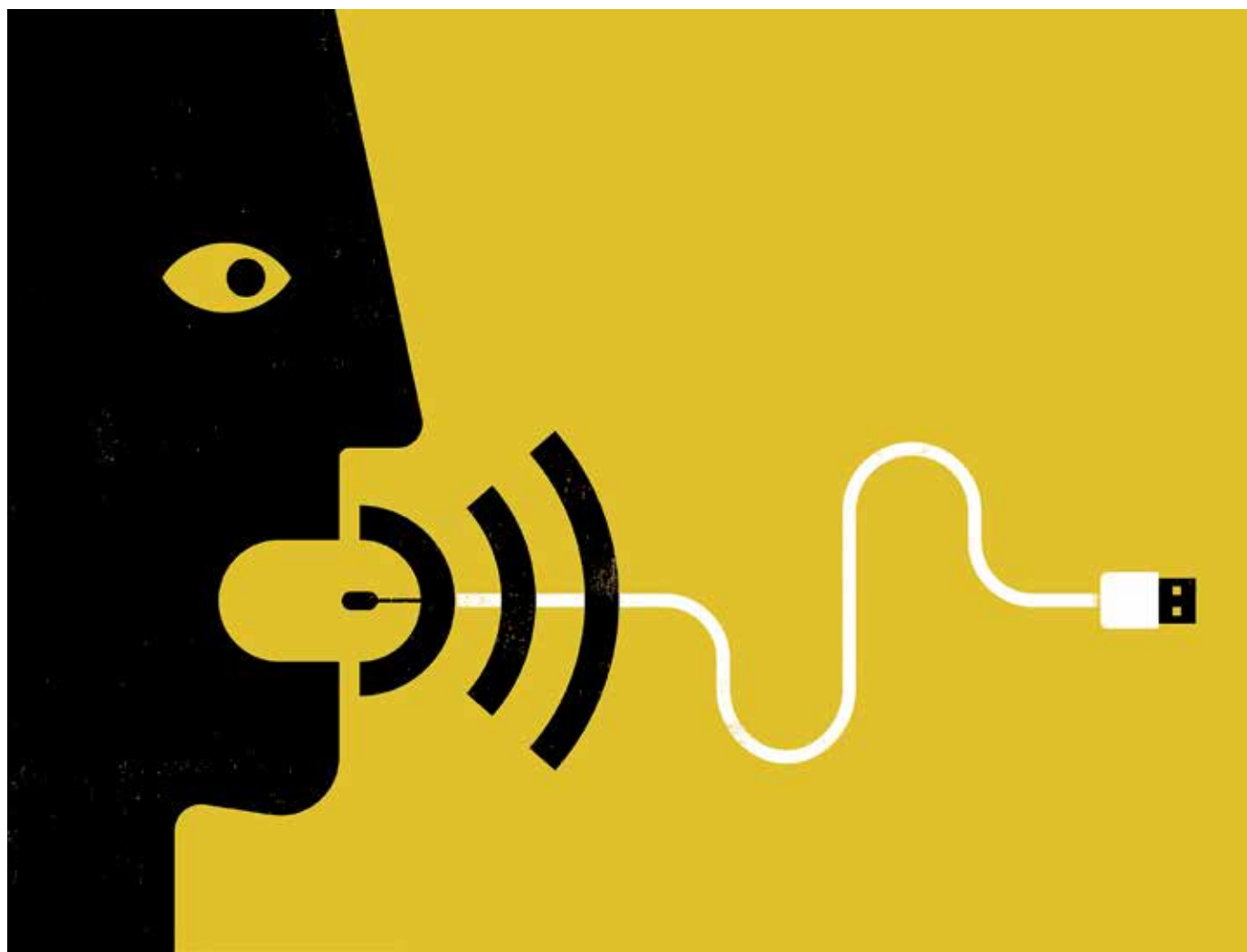# WRITING CODE OUT LOUD

Programmers turn to voice-command tools to give their hands a rest.

ILLUSTRATION BY THE PROJECT TWINS

BY JULIAN NOWOGRODZKI

Debilitating hand pain is always bad news, but Harold Pimentel's was especially unwelcome. As a computational-biology PhD student, his work involved constant typing — and he was born with only one arm. "My adviser jokingly said, 'Can't you do this by voice?'" he recalls. Three years later, as a computational-genomics postdoc at Stanford University in California, he does just that.

Pimentel had cubital tunnel syndrome caused by repetitive strain injury (RSI). The syndrome occurs when the ulnar nerve, which travels down the outer edge of the arm, becomes pinched at the elbow, causing numbness, pain and loss of fine motor control in the hands and fingers. RSI can derail the careers of computational biologists and other scientists who code. Now, a small but growing community has developed a workaround: coding by voice command. It takes at least a month

of difficult, sometimes frustrating, training to get set up, but coding by voice helps these programmers to keep doing their jobs or continue their studies. And they say that there are unexpected advantages.

YOUTUBE INSPIRATION

Voice coding underlies a wide variety of science — any researcher who writes code could use it. Matthew Solomonson, a software engineer at the Broad Institute of MIT and ▶

▶ Harvard in Cambridge, Massachusetts, uses it to build web applications such as the Genome Aggregation Database (gnomAD), which is used to explore genomic data. "These applications share data from some of the largest sequencing studies in the world," he says.

Naomi Saphra, a PhD student in language cognition and computing at the University of Edinburgh, UK, has small-fibre neuropathy, the cause of which is unknown. It is a permanent condition of the nerves that connect the brain to the hands and feet, and causes the nerves to transmit pain in response to sensations that are not usually painful. She uses her code to explore the training process of neural language models. And Pimentel studies how retention of the non-coding sections of RNA determines tissue specificity and disease susceptibility.

Similarly to many other scientists, Pimentel and Saphra realized that voice coding was possible thanks to a video of Tavis Rudd, now director of technology at web-development firm Unbounce, demonstrating the process live at the PyCon 2013 conference for users of the Python programming language.

In that video, Rudd describes his struggle with RSI — the result of constant coding in the emacs text editor, a condition he calls 'emacs pinkie' — and his strategy for overcoming it. He developed a solution through months of painstaking work, and he calls it a "three-headed beast" because it runs three operating systems from one laptop. In front of the crowd at the conference, he used his method to dictate code instructing his laptop to read aloud a snippet of Monty Python's Dead Parrot sketch.

"It was pretty inspirational," Pimentel says. But the process also was "super buggy", Pimentel reports, and lacked an active user community to help fix the glitches. He began looking for alternatives, as did Saphra.

Coding by voice command requires two kinds of software: a speech-recognition engine and a platform for voice coding. Dragon from Nuance, a speech-recognition software developer in Burlington, Massachusetts, is an advanced engine and is widely used for programming by voice, with Windows and Mac versions available. Windows also has its own built-in speech recognition system. On the platform side, VoiceCode by Ben Meyer and Talon by Ryan Hileman (both are for Mac OS only) are popular.

Two other platforms for voice programming are Caster and Aenea, the latter of which runs on Linux. Both are free and open source, and enable voice-programming functionality in Dragonfly, which is an open-source Python framework that links actions with voice commands detected by a speech-recognition engine. Saphra tried Dragonfly, but found that setup required more use of her hands than she could tolerate.

All of these platforms for voice command work independently of coding language and text editor, and so can also be used for tasks outside programming. Pimentel, for instance, uses voice recognition to write e-mails, which he finds easier, faster and more natural than typing.

STACCATO BURSTS

To the untrained ear, coding by voice command sounds like staccato bursts of a secret language. Rudd's video is full of terms like 'slap' (hit return), 'sup' (search up) and 'mara' (mark paragraph).

Unlike virtual personal assistants such as Apple's Siri or Google's Alexa, VoiceCode and Talon don't do natural-language processing, so spoken instructions have to precisely match the commands that the system already knows. But both platforms use continuous command recognition, so users needn't pause between commands, as Siri and Alexa require.

VoiceCode commands typically use words not in the English language, because if you use an English word as a command, such as 'return', it means you can never type out that word. By contrast, Talon, Aenea and Caster feature dynamic grammar, a tool that constantly updates which words the software can recognize on the basis of which applications are open. This means users can give English words as commands without causing confusion.

In addition to voice recognition, Talon can also replace a computer mouse with eye tracking, which requires a Tobii 4c eye tracker (US$150). Other eye-mousing systems generally require both the eye tracker and head-tracking hardware, such as the TrackIR from NaturalPoint. "I want to make fully hands-free use of every part of a desktop computer a thing," says Hileman. Other mouse replacements also exist; Pimentel uses one called SmartNav.

Voice command requires at least a decent headset or microphone. Many users choose a unidirectional microphone so that others can talk to them while they are dictating code. One such mic, a cardioid mic, requires special equipment to supply power, and hardware costs can reach $400, says Pimentel.

The software can cost several hundred dollars too. The speech-recognition engine Dragon Professional costs $300, as does VoiceCode. Caster and Aenea are free and open source. Talon is available for free, but requires a separate speech-recognition engine. A beta version of Talon that includes a built-in speech-recognition engine is currently available to Hileman's Patreon supporters for $15 per month. "This kind of tech needs to be as free and widespread as possible, because I feel like we're sitting on an RSI epidemic in progress and nobody is talking about it," says Hileman. But, he adds, a "huge goal" of his is to convince lots of people who are not yet experiencing problems".

IT TAKES A VILLAGE

Whether or not users have RSI, it can be difficult and frustrating to start programming by voice. It took a month and a half for Pimentel to get up to speed, he says, and there were days when he was ready to throw in the towel. He printed out 40 pages of commands and forced himself to look at them until he learnt them. Saphra needed two months of coding, a little every day, before she felt that it was a "perfectly enjoyable experience and I could see myself doing this for a living".

After the initial learning curve, users often create custom prompts for commonly used commands as the need arises. Saphra has written prompts for creating fractions in the mathematical-typesetting system LaTeX.

Users often share their configuration files and set-up details on sites such as GitHub and Slack. VoiceCode's Slack channel has more than 250 users, of whom perhaps 40 are active, Pimentel estimates; the Talon Slack has more than 100, and some users are in both. "The communities are really important for both of these," says Saphra. "These are not tools you can use without being deeply involved with the community."

"These are not tools you can use without being deeply involved with the community."

Pimentel and Saphra are both from the United States, as are most of the coders they know who use voice recognition. The software is usually worse at interpreting speakers who have an accent other than standard American. According to an analysis (see go.nature.com/2ffx78z) by data scientist Rachael Tatman, at least some tools make more errors with women's voices. "If you don't have a mainstream American accent, or if you're a woman, then it's going to be a much more painful process," says Saphra. "But even then, it's not so bad."

There are other downsides. Pimentel has throat problems and has to take frequent breaks. "I drink so much frickin' water," he says. He's looking into voice training to learn to put less strain on his voice.

Pimentel misses being able to work in a quiet library, while Saphra misses being able to make more noise. "I used to listen to music or sing to myself while I coded. Or I would just curse. I can't do that anymore," she says.

But they appreciate the benefits, too. "I've often thought that if I woke up one day and my hands were miraculously perfectly fine, that I would continue to dictate code," Saphra says, only typing if something was tricky to dictate. "I think a lot of people could benefit from it." Voice command can even be relaxing, Saphra notes: she can put her feet up instead of hunching over a keyboard.

Probably the greatest extra benefit, though, is more intangible. "Now I really think carefully before I say anything or execute anything," says Pimentel, "and my code seems to have fewer bugs." Saphra adds: "I feel a mastery of my own tools in a way that I was never motivated to gain before this." ∎

Julian Nowogrodzki *is a freelance writer based near Boston, Massachusetts.*