AI interprets marmosets' trills, chirps and peeps

BY **BAHAR GHOLIPOUR**

14 NOVEMBER 2017

New artificial intelligence software can decode conversations between small monkeys called marmosets. Researchers reported the findings today at the **2017 Society for Neuroscience annual meeting** in Washington D.C.

Marmosets are social animals that live in groups. Their vocabulary includes 10 to 15 calls — twitters, trills, chirps, 'phees' and various kinds of peeps, each with its own meaning. Studies suggest that like human babies, baby marmosets learn to communicate by hearing other marmosets talk to them[1,2].

The monkeys' human-like communication system has made them **popular among scientists** who study language, social communication or vocal production. Marmosets carrying autism-linked mutations are also good models for **altered social behavior** in the condition.

Visual representations of marmoset calls, called spectrograms, also appear as distinct patterns of frequencies. But it is tedious to manually extract the different calls from hours of recordings of screeching monkeys, cage noises and people talking in the background.

"Either you hire a bunch of graduate students to sit through and listen to all of it and classify it or you come up with a way for a computer to do it," says Samvaran Sharma, an engineer at the McGovern Institute at the Massachusetts Institute of Technology in Cambridge, Massachusetts, who presented the work. Sharma developed the software with colleagues in **Robert Desimone**'s laboratory at MIT.

The interlocutor:

The new program first separates the audio from background noise and detects which segments of

the recording contain a monkey call. Spectrograms of those segments are then turned into black and white images. So, for example, a chirp looks like a series of thin vertical white brush strokes, while a phee looks like a thick, slightly curved dash.

Sound bites: An algorithm translates frequency patterns from marmoset calls into pictures.

These letter-like images are then passed on to an artificial neural network for classification.

Neural networks are algorithms modeled after the brain that excel at finding patterns. To learn the calls, the neural network takes a bunch of examples — say, the images created to represent a phee. It works out what core characteristics define a phee and then uses that pattern to figure out whether a new picture is a phee or something else.

The result is a nicely classified string of marmoset words that researchers can read. From there, they can piece together what the monkeys are saying to each other.

To test the program, the team used samples from hours of audio from small groups of marmosets as they roamed around in their cage interacting with each other. Wireless mics that the monkeys wore on their collars recorded the audio.

The software sifts monkey talk from background noise with 80 percent accuracy and correctly identifies which call the monkeys make more than 90 percent of the time.

The team is working to improve the program's accuracy. They plan release the code as an out-of-the-box program that requires no expertise to use, Sharma says.

*For more reports from the 2017 Society for Neuroscience annual meeting, please **click here**.*

REFERENCES:

1. Takahashi D.Y. *et al*. Science **349**, 734-738 (2015) **PubMed**
2. Takahashi D.Y. *et al.* Philos. Trans. R Soc. Lond. B. Biol. Sci. **371** (2016) **PubMed**