pit autonomous drones against each other[10] — a development that will keep pushing the boundaries of this widely relevant technology.

Guido C. H. E. de Croon is in the Faculty of Aerospace Engineering, Delft University of Technology, 2629 HS Delft, the Netherlands.
e-mail: g.c.h.e.decroon@tudelft.nl

1. Campbell, M., Hoane, A. J. Jr & Hsu, F. H. *Artif. Intell.* **134**, 57–83 (2002).
2. Silver, D. *et al. Nature* **550**, 354–359 (2017).
3. Wurman, P. R. *et al. Nature* **602**, 223–228 (2022).
4. Kaufmann, E. *et al. Nature* **620**, 982–987 (2023).
5. Beard, R. W. *Quadrotor Dynamics and Control* (Brigham Young Univ., 2008).
6. O'Connell, M. *et al. Sci. Robot.* **7**, eabm6597 (2022).
7. Salvato, E., Fenu, G., Medvet, E. & Pellegrino, F. A. *IEEE Access* **9**, 153171–153187 (2021).
8. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 2018).
9. Scheper, K. Y. W. & de Croon, G. C. H. E. In *Proc. 14th International Conference on Simulation of Adaptive Behavior* (eds Tuci, E., Giagkos, A., Wilson, M. & Hallam, J.) 280–292 (Springer, 2016).
10. De Wagter, C., Paredes-Vallés, F., Sheth, N. & de Croon, G. C. H. E. *Field Robot.* **2**, 1263–1290 (2022).

The author declares no competing interests.

Clinical neuroscience

Speech-enabling brain implants pass milestones

Nick F. Ramsey & Nathan E. Crone

Two brain–computer interfaces have been developed that bring unprecedented capabilities for translating brain signals into sentences — at speeds close to that of normal speech, and with vocabularies exceeding 1,000 words. **See p.1031 & p.1037**

There is an urgent need to help people with neurological conditions that deprive them of the universal human need to communicate. Two articles published in *Nature* demonstrate that individuals who are unable to speak as a result of severe paralysis could potentially use implantable brain–computer interfaces (BCIs) to communicate at rates much greater than those typically achievable with alternative communication options. Willett *et al.*[1] (page 1031) report a device that records brain activity using electrodes that penetrate the brain's cortex, whereas Metzger and colleagues' device[2] (page 1037) uses electrodes placed on the cortical surface. These studies signal a turning point in the development of BCI technology that aims to restore communication for people with severe paralysis.

Various neurological disorders paralyse muscles crucial to speech and limb function while sparing cognitive functions, potentially resulting in locked-in syndrome — in which individuals can no longer initiate communication and can respond to queries only with eye blinks or minimal movements. A diverse range of systems, known as alternative and augmentative communication technologies, are available to help people with locked-in syndrome to communicate, but these require effort and are much slower (achieving, typically, just a few words per minute) than normal speech (about 150 words per minute). BCIs have the potential to solve these problems.

The first demonstration that a subject could be trained to increase the activity of single neurons, and thereby to exert a wilful action, was published in 1969, for a rhesus macaque (*Macaca mulatta*)[3]. Experiments in humans began[4] in the late 1990s, when an electrode was connected to neurons in a person with locked-in syndrome caused by motor neuron disease (amyotrophic lateral sclerosis, or ALS), a neurodegenerative disease. This was followed in 2006 by a study[5] in which arrays of millimetre-scale electrodes (known as microelectrodes) were implanted into the brain of a person with a spinal cord injury. This microelectrode array (MEA) recorded the activity of several hundred neurons in the motor cortex, the brain region responsible for the control of voluntary movements, and thereby controlled a robotic arm[5]. MEAs have since been used to enable communication, for instance by decoding handwriting attempts[6].

The complementary technique of electroencephalography (EEG) — in which electrodes are placed along the scalp to record electrical activity in the brain — has been used since 1999 (ref. 7) to help people with paralysis to communicate by controlling custom spelling software[8]. Around the same time, it was discovered that small disc-shaped electrodes (2–3 millimetres in diameter) placed on the surface of the brain could acquire much higher-quality signals than could be obtained using scalp electrodes[9]. This method for recording brain activity is known as electrocorticography (ECoG).

In the early 2000s, ECoG electrodes were used in people undergoing surgery for drug-resistant epilepsy, to record brain signals associated with speech and body movements[10]. This eventually led to the development of the first fully embedded ECoG device, which enabled a person with locked-in-syndrome to use a typing program at home[11]. To date, about 50 people with varying degrees of paralysis have been implanted with BCIs for communication, most of whom use MEAs.

Metzger *et al.* now present findings involving a paralysed participant who, 17 years before she enlisted for the study, experienced a brainstem stroke that made her speech unintelligible. The authors' BCI system incorporates a silicon sheet embedded with 253 ECoG electrodes, each of which record the average activity of many thousands of neurons (Fig. 1a). The device was surgically implanted over the left 'face area' of the sensorimotor cortex — the part of the brain that serves oral and facial muscles, including the vocal tract. The study builds on previous reports of ECoG recordings, including a similar BCI that was implanted in another person who had had a brainstem stroke[12].

Brain-to-text decoding was achieved by the combination of two systems: a recurrent neural network (RNN, a type of artificial neural network), which ran algorithms that decipher brain activity associated with movements of articulators (parts of the vocal tract); followed by a language model that composed sentences at a rate of 78 words per minute (albeit with a 25.5% word error rate) from a set of 1,024 words. Alternatively, brain signals were translated directly to synthesized speech, at a word error rate of 54.4% for the 1,024-word vocabulary; the error rate decreased for smaller vocabularies (8.2% for a 119-word vocabulary). The BCI also decoded attempted facial expressions, which it reproduced using a digital avatar, thereby providing visual feedback to the text or speech that greatly enriches the participant's ability to communicate. Overall, the device offers substantial improvements in vocabulary size, speed of communication and versatility of speech decoding compared with previously reported ECoG BCIs.

Willett *et al.*[1] used two MEAs (containing a total of 128 electrodes) to record from small patches of the left sensorimotor face area in a participant who was unable to speak intelligibly owing to ALS (Fig. 1b). As in Metzger and colleagues' device, RNNs and language models were used to translate brain signals into text and were trained and tested on vocabularies of different sizes. Using the device, the participant was able to communicate at an average rate of 62 words per minute, with a word error rate of 23.8% for a 125,000-word vocabulary and 9.1% for a 50-word vocabulary.

The RNN was trained using recordings of neural activity collected when the participant attempted to speak 260–480 sentences presented on a monitor — the overall process

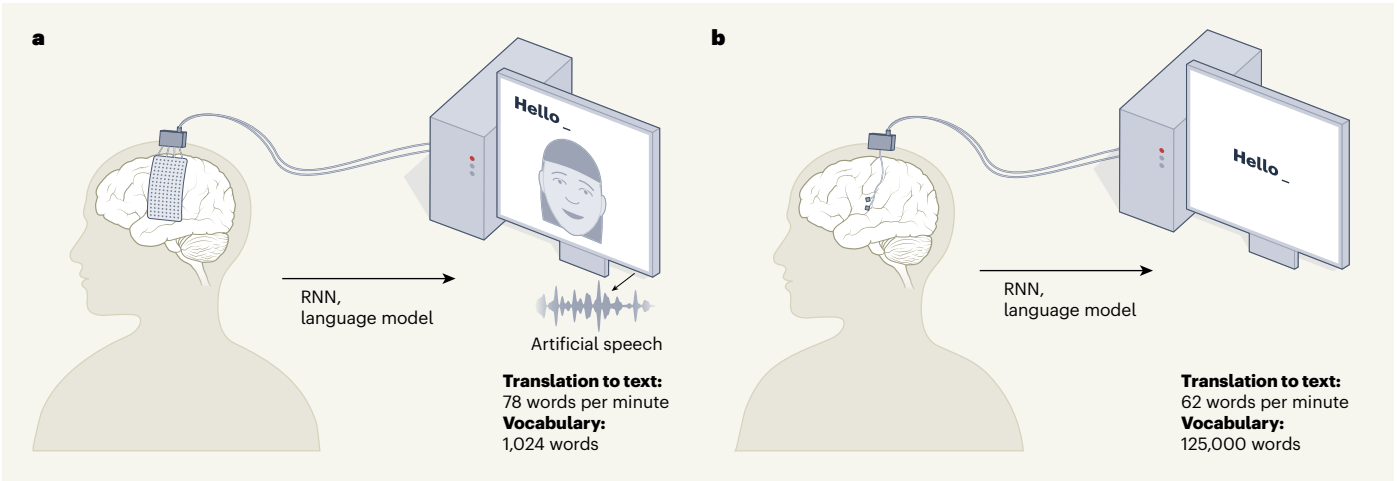

Figure 1 | Advanced techniques for translating thoughts into speech. Two brain–computer interfaces (BCIs) have been tested in individuals who cannot speak intelligibly as a result of paralysis. **a**, Metzger and colleagues' device[2] uses electrodes placed on the surface of a wide area of the brain's cortex to record brain activity, which is translated into speech or text using a recurrent neural network (RNN, a type of artificial neural network). A language model is used to reduce errors in the composed sentences. The BCI also translates brain activity into facial expressions, which are represented using an avatar. **b**, Willett and colleagues' device[1] uses arrays of microelectrodes implanted into the cortex, and records signals from a relatively small number of neurons. The brain activity is converted into text using an RNN and a language model. The two studies demonstrate that BCIs can translate neural activity into speech at speeds approaching those of normal speech (about 150 words per minute), and that uses large vocabularies.

took an average of 140 minutes every day, for 8 days. Willett and colleagues present analyses suggesting that this daily training could be reduced considerably without much loss of performance. Importantly, the authors observed that neural activity recorded from a brain region (called Broca's area) widely thought to be crucial for speech production could not be decoded — raising questions about whether this area contains useful information for speech decoding.

The two reports constitute crucial proof of concept that communication can be restored using implantable BCIs, but several issues require further investigation to allow for more widespread use. First, the speech models used in both studies were trained and tested using mimed speech of participants who had residual, albeit weak, articulatory movements. More studies are now needed to show efficacy in participants who lack residual movements, as occurs in locked-in syndrome (including late-stage ALS). Another issue is that, for both devices, high-bandwidth recordings were taken from hundreds of electrodes, which had to be connected to external amplifiers through a 'pedestal' that penetrates the skin, which is cosmetically unappealing. Fully implantable, wireless BCIs will need to be developed that replicate or surpass the performance reported in these studies.

Furthermore, highly skilled researchers were actively involved in the operation of the reported BCIs, which remain too complicated for caregivers to operate in home settings without extensive training and maintenance. Similarly effective BCI systems that operate with minimal or no researcher intervention will be needed in the future. This will require extensive development and testing in clinical populations, using user-centred design principles. It is also unclear whether the user's perception of other people's speech will cause errors in brain-to-text decoding, given increasing evidence that speech perception, in addition to speech production, activates the sensorimotor cortex[13,14].

Finally, it remains to be seen which BCI approach — MEAs or ECoG — will best serve the needs of users in terms of safety and long-term efficacy in real-life applications. MEAs capture rich functional information from a small cortical area, but the signals tend to be unstable and require frequent updating of speech-decoding models. Moreover, the longevity of MEAs might be limited by degradation of the electrode materials and tissue encapsulation of the devices[15]. ECoG electrodes need to be implanted over a larger area than do MEAs, but the ECoG electrodes are external to the cortical tissue and usually deliver excellent signal quality for many years[16], although they can elicit superficial tissue reactions[17].

In the meantime, the two BCIs represent a great advance in neuroscientific and neuroengineering research, and show great promise in boosting the quality of life of individuals who have lost their voice as a result of paralysing neurological injuries and diseases. Even a basic BCI implant that allows the user to select letters or icons in assistive-technology software provides them with considerable benefits and satisfaction in daily life[18]. Advanced BCI systems that enable communication, such as those discussed here, can be expected to have an even greater impact.

Nick F. Ramsey is in the UMC Utrecht Brain Center, Department of Neurology and Neurosurgery, University Medical Center Utrecht, 3584 CX Utrecht, the Netherlands. **Nathan E. Crone** is in the Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, Maryland 21218, USA. e-mails: n.f.ramsey@umcutrecht.nl; ncrone@jhmi.edu

1. Willett, F. R. *et al. Nature* **620**, 1031–1036 (2023).
2. Metzger, S. L. *et al. Nature* **620**, 1037–1046 (2023).
3. Fetz, E. E. *Science* **163**, 955–958 (1969).
4. Kennedy, P. R. & Bakay, R. A. *NeuroReport* **9**, 1707–1711 (1998).
5. Hochberg, L. R. *et al. Nature* **442**, 164–171 (2006).
6. Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M. & Shenoy, K. V. *Nature* **593**, 249–254 (2021).
7. Kübler, A. *et al. Exp. Brain Res.* **124**, 223–232 (1999).
8. Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G. & Vaughan, T. M. *Clin. Neurophysiol.* **113**, 767–791 (2002).
9. Crone, N. E., Miglioretti, D. L., Gordon, B. & Lesser, R. P. *Brain* **121**, 2301–2315 (1998).
10. Leuthardt, E. C., Schalk, G., Wolpaw, J. R., Ojemann, J. G. & Moran, D. W. *J. Neural Eng.* **1**, 63–71 (2004).
11. Vansteensel, M. J. *et al. N. Engl. J. Med.* **375**, 2060–2066 (2016).
12. Moses, D. A. *et al. N. Engl. J. Med.* **385**, 217–227 (2021).
13. Wilson, S., Saygin, A., Sereno, M. I. & Iacoboni, M. *Nature Neurosci.* **7**, 701–702 (2004).
14. Berezutskaya, J., Baratin, C., Freudenburg, Z. V. & Ramsey, N. F. *Hum. Brain Mapp.* **41**, 4587–4609 (2020).
15. Woeppel, K. *et al. Front. Bioeng. Biotechnol.* **9**, 759711 (2021).
16. Sillay, K. A. *et al. Brain Stimul.* **6**, 718–726 (2013).
17. Degenhart, A. D. *et al. J. Neural Eng.* **13**, 046019 (2016).
18. Pels, E. G. M. *et al. Clin. Neurophysiol.* **130**, 1798–1803 (2019).

The authors declare no competing interests.
This article was published online on 23 August 2023.