ORIGINAL ARTICLE

# An Accurate and Rapidly Calibrating Speech Neuroprosthesis

Nicholas S. Card, Ph.D., Maitreyee Wairagkar, Ph.D., Carrina Iacobacci, B.S., Xianda Hou, M.S., Tyler Singer-Clark, B.S., Francis R. Willett, Ph.D., Erin M. Kunz, M.S., Chaofei Fan, M.S., Maryam Vahdati Nia, M.S., Darrel R. Deo, Ph.D., Aparna Srinivasan, M.S., Eun Young Choi, Ph.D., Matthew F. Glasser, M.D., Ph.D., Leigh R. Hochberg, M.D., Ph.D., Jaimie M. Henderson, M.D., Kiarash Shahlaie, M.D., Ph.D., Sergey D. Stavisky, Ph.D., and David M. Brandman, M.D., Ph.D.

## ABSTRACT

### BACKGROUND

Brain–computer interfaces can enable communication for people with paralysis by transforming cortical activity associated with attempted speech into text on a computer screen. Communication with brain–computer interfaces has been restricted by extensive training requirements and limited accuracy.

### METHODS

A 45-year-old man with amyotrophic lateral sclerosis (ALS) with tetraparesis and severe dysarthria underwent surgical implantation of four microelectrode arrays into his left ventral precentral gyrus 5 years after the onset of the illness; these arrays recorded neural activity from 256 intracortical electrodes. We report the results of decoding his cortical neural activity as he attempted to speak in both prompted and unstructured conversational contexts. Decoded words were displayed on a screen and then vocalized with the use of text-to-speech software designed to sound like his pre-ALS voice.

### RESULTS

On the first day of use (25 days after surgery), the neuroprosthesis achieved 99.6% accuracy with a 50-word vocabulary. Calibration of the neuroprosthesis required 30 minutes of cortical recordings while the participant attempted to speak, followed by subsequent processing. On the second day, after 1.4 additional hours of system training, the neuroprosthesis achieved 90.2% accuracy using a 125,000-word vocabulary. With further training data, the neuroprosthesis sustained 97.5% accuracy over a period of 8.4 months after surgical implantation, and the participant used it to communicate in self-paced conversations at a rate of approximately 32 words per minute for more than 248 cumulative hours.

### CONCLUSIONS

In a person with ALS and severe dysarthria, an intracortical speech neuroprosthesis reached a level of performance suitable to restore conversational communication after brief training. (Funded by the Office of the Assistant Secretary of Defense for Health Affairs and others; BrainGate2 ClinicalTrials.gov number, NCT00912041.)

From the Departments of Neurological Surgery (N.S.C., M.W., C.I., X.H., T.S.-C., M.V.N., A.S., K.S., S.D.S., D.M.B.), Computer Science (X.H., M.V.N.), and Biomedical Engineering (T.S.-C., A.S.), University of California, Davis, Davis, and the Departments of Neurosurgery (D.R.D., E.Y.C., J.M.H.), Electrical Engineering (E.M.K.), and Computer Science (C.F.), the Wu Tsai Neurosciences Institute (E.M.K., J.M.H.), the Howard Hughes Medical Institute (F.R.W.), and Bio-X (J.M.H.), Stanford University, Stanford — both in California; the Departments of Radiology and Neuroscience, Washington University School of Medicine, Saint Louis (M.F.G.); the School of Engineering and Carney Institute for Brain Sciences, Brown University (L.R.H.), and the Center for Neurorestoration and Neurotechnology, Department of Veterans Affairs Office of Rehabilitation Research and Development, VA Providence Healthcare (L.R.H.) — both in Providence, RI; and the Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston (L.R.H.). Dr. Card can be contacted at nscard@ucdavis.edu, Dr. Stavisky at sstavisky@ucdavis.edu, and Dr. Brandman at dmbrandman@ucdavis.edu.

Drs. Stavisky and Brandman contributed equally to this article.

N Engl J Med 2024;391:609-18.
DOI: 10.1056/NEJMoa2314132

Communication is a priority for people with dysarthria from neurologic disorders such as stroke and amyotrophic lateral sclerosis (ALS).[1] People with diseases that impair communication have an increased risk of isolation, depression, and decreased quality of life[2,3]; losing communication may determine whether a person will pursue or withdraw life-sustaining care in advanced ALS.[4] Although augmentative and assistive communication technologies such as eye trackers (also called eye-gaze–tracking devices) or head trackers are available, they have low information-transfer rates and become increasingly difficult to use as patients lose voluntary muscle control.[5]

Brain–computer interfaces are a promising type of communication technology that can directly decode the user's intended speech from neural signals.[6] Efforts to develop a speech neuroprosthesis are built largely on studies involving data that are retrospectively analyzed from able-bodied speakers who undergo electrophysiological monitoring for clinical purposes.[7-16] Several groups have performed studies of real-time brain–computer interfaces for the restoration of lost speech with electrocorticographic arrays implanted on the cortical surface[17-20] (one of which was reported in the *Journal*[17]) or intracortical multielectrode arrays.[21] Two recent studies have established "brain-to-text" speech performance[19,21] by decoding the cortical neural signals generated during attempted speech into phonemes (the building blocks of words) and assembling these phonemes into words or sentences displayed on a computer screen. In these studies, the median word error rate (i.e., the percentage of words that were decoded incorrectly) was 25.5% with a 1024-word vocabulary[19] and 23.8% with a 125,000-word vocabulary[21]; approximately 17 hours of recording were required in order to collect training data that were sufficient to obtain that level of performance.

*An audio recording of the participant's speech is available at NEJM.org*

We report an intracortical speech neuroprosthesis with low training-data requirements that ultimately provided access to a 125,000-word vocabulary in a person with advanced ALS and severe dysarthria. The neuroprosthesis achieved high accuracy with useful function beginning on the first day of use, 25 days after implantation.

## Methods

### Study Participant

A 45-year-old left-handed man with ALS had symptoms that had begun 5 years before enrollment into this study. At the time of enrollment, he was nonambulatory; was dependent on others for controlling his electric wheelchair, dressing, eating, and hygiene; had severe dysarthria; and had an ALS Functional Rating Scale–Revised score of 23 (scores range from 0 to 48, with higher scores indicating better function). For 8 consecutive months after surgical placement of the recording arrays, he has continued to have a modified Mini–Mental State Examination score of 27 (scores range from 0 to 27, with higher scores being consistent with better cognitive function).

At the time of this report, he retained eye and neck movements but had limited orofacial movement and a mixed upper- and lower-motor neuron dysarthria that resulted in monotone, low-volume, nasal speech. When his speech was being listened to by people who were not his regular care partner, it was unintelligible (listen to audio): his oral motor tasks on the Frenchay Dysarthria Assessment–Second Edition (a measure of several speech behaviors) received an E rating, indicating profound dysarthria (ratings range from A to E, with A representing normal function and E no function). When speaking to expert listeners, he communicated at a mean (±SD) rate of 6.8±5.6 correct words per minute (the rate of conversational English is approximately 160 words per minute[22]). His mean typing speed when he used a gyroscopic head mouse (Zono 2, Quha) was 6.3±1.3 correct words per minute (Fig. S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). The severity of dysarthria remained stable during the period of this report, including the immediate postoperative period. At the time of this report, he received noninvasive respiratory support at night and did not have a tracheostomy. Additional details about the participant are provided in Section S1.01 of the Supplementary Appendix.

Informed consent was obtained from the participant with the use of an informed-consent form that was approved by the institutional review board at Mass General Brigham; the informed-consent form was also approved locally by the institutional review board at the University of California, Davis. The site-responsible principal investigator (the last author), the co–senior author (the second-to-last author), and the sponsor-investigator for the clinical study (the fifth-to-last author) vouch for the accuracy and completeness of the data and for the fidelity of the study to the protocol, which is available at NEJM.org. Neither Blackrock Neurotech (the manufacturer of the arrays and signal-processing system) nor any other commercial entity was involved in data collection or reporting in this study or had oversight regarding the decision to publish these results. There were no study-related agreements of any kind between the authors and a commercial entity. The devices used were purchased for research use and were not provided by a commercial entity.

A total of 19 participants have participated in the BrainGate and ongoing BrainGate2 clinical studies, which historically focused on decoding attempted arm and hand movements from related areas of cortex. After the recent evolution of the trial to include recording from speech areas of cortex, results for a speech neuroprosthesis in one previous participant have been reported[21]; that participant had only two arrays implanted in the precentral gyrus (and two in the inferior frontal gyrus) rather than the four arrays implanted in the precentral gyrus in the current participant.[21]

#### SURGICAL IMPLANTATION

We implanted four microelectrode arrays (NeuroPort Array, Blackrock Neurotech) into the left precentral gyrus, an important cortical region for coordinating motor activities related to speech.[17,19,21] Each microelectrode array measures 3.2 mm by 3.2 mm and has 64 electrodes arranged in an 8-by-8 grid inserted 1.5 mm into the cortex with the use of a specialized high-speed pneumatic inserter. Each electrode has one recording site approximately 50 $\mu$m in diameter and is designed to record from one or a small number of cortical neurons. The arrays were implanted through a 5 cm by 5 cm craniotomy on the left side under general anesthesia. Care was taken to avoid placing the microelectrode arrays through large vessels on the cortical surface that were identified by visual inspection. Two arrays are connected to one percutaneous connector (“pedestal”) designed to transmit the neural recordings to external computers. Two percutaneous pedestals, each secured to the skull with titanium screws, provide for recording from a total of 256 sites. Reference wires are placed in the subdural and epidural spaces. The pedestals are connected by detachable connectors that use HDMI cables to transmit data to computers (Fig. 1A). These computers sit on a wheeled cart and are connected to standard electrical wall outlets.

The surgical implantation was performed in July 2023; no serious adverse events occurred, and the participant was discharged on postoperative day 3. Nonserious adverse events, including incisional pain and a transient increased frequency of the muscle spasms in the limbs from spasticity that he had had for months before implantation, are listed in Section S1.01. From the initial incision to closure, the operation lasted 5 hours. We began collecting data in August 2023, at 25 days after surgery.

#### RECORDING ARRAY LOCATIONS AND DECODING CONTRIBUTIONS

Before implanting arrays in the precentral gyrus, we used magnetic resonance imaging (MRI) to identify the central sulcus and used functional MRI (fMRI) and standard clinical fMRI tasks (sentence completion, silent word generation, silent verb generation, and object naming) to confirm that the participant was left-hemisphere language dominant. We refined the implantation targets using the Human Connectome Project multimodal MRI-derived cortical parcellation precisely mapped to the participant’s brain[23] (Fig. 1B and Fig. S2 and Section S1.02; Fig. S11 shows the estimated locations on the Montreal Neurological Institute template brain). We targeted language-related area 55b[24] (an area implicated in phonologic representation) and three areas in the ventral precentral gyrus associated with speech production: the dorsal and ventral aspects of the ventral premotor cortex and the primary motor cortex (Brodmann area 4). Our choice to target the speech motor cortex was informed by our previous study involving another (aforementioned) participant, in which two arrays in the ventral premotor cortex were shown to provide informative signals for speech decoding.[21]

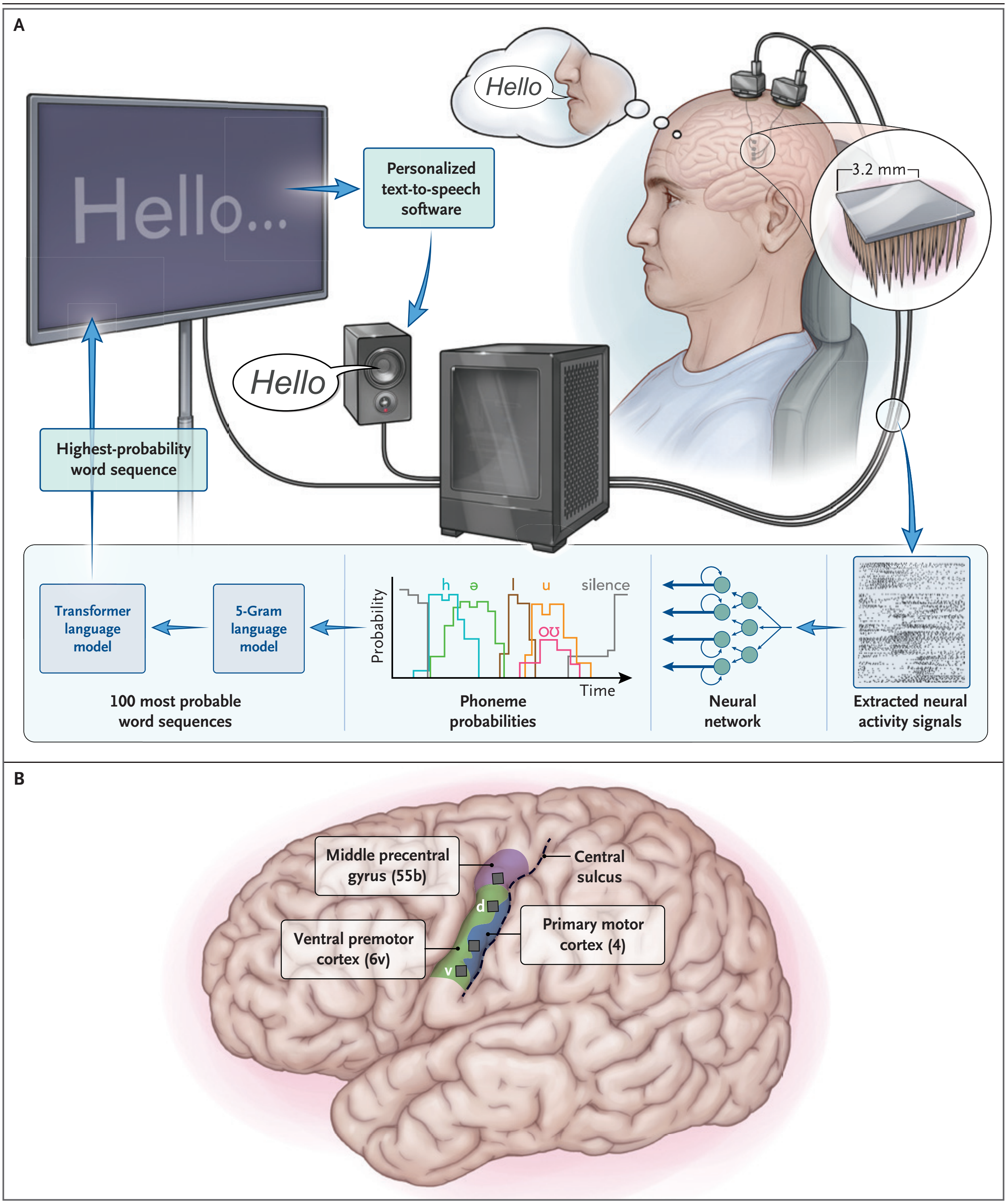

A
Hello
Hello...
Personalized text-to-speech software
3.2 mm
Hello
Highest-probability word sequence
Transformer language model
5-Gram language model
100 most probable word sequences
Probability
h
ə
l
u
oʊ
silence
Time
Phoneme probabilities
Neural network
Extracted neural activity signals
B
Middle precentral gyrus (55b)
Central sulcus
d
Ventral premotor cortex (6v)
Primary motor cortex (4)
v

**Figure 1 (facing page). Electrode Locations and Speech-Decoding Setup.**

Panel A shows the brain-to-text speech neuroprosthesis. Electrical activity is measured with the use of four 64-electrode arrays and processed to extract neural activity (see Section S1.04). Machine-learning techniques, incorporating an artificial neural network, are used to decode the extracted neural activity signals into an English phoneme every 80 msec (see also Section S5). Panel B shows the approximate microelectrode array locations (gray squares) superimposed on a three-dimensional reconstruction of the participant's brain. Colored regions correspond to cortical areas[22] aligned to the participant's brain with the use of the Human Connectome Project MRI protocol scans before implantation.

### Real-Time Acquisition and Processing of Cortical Recordings

A signal-processing system (NeuroPort System, Blackrock Neurotech) was used to acquire signals from the two connector pedestals (Fig. S3) and send them to a series of commercially available computers running our publicly available software[25] (Section S1.5) for real-time signal processing (Section S1.4) and decoding (Sections S2 and S3). A researcher connected and disconnected the pedestals at the start and end of each session.

### Speech Task Designs

We collected data in 84 sessions over 32 weeks (Section S1.06 and Table S2) in the participant's home. No more than one session was held on any study day. Each study session consisted of a series of task blocks, lasting approximately 5 to 30 minutes, during which the participant used the neuroprosthesis. Between blocks, the participant would take breaks. During each block, the participant used the system in two different ways: an instructed-delay copy task (Videos 1 and 2 and Section S1.07) and a participant-paced conversation mode (Videos 4 and 5 and Section S1.08). The instructed-delay task consisted of words being shown on a computer screen, with the participant attempting to say the words after a visual or audio cue.[21] The conversation mode involved the participant attempting to say whatever he wanted (although the computer outputs were limited to a 125,000-word dictionary) in an unstructured conversational context. During both tasks, speech decoding occurred in real time; as the participant attempted to speak, the cortical activity at the four microelectrode arrays was recorded and decoded, and the predicted words were shown on the screen. Completed sentences were read aloud by a computer program and, in later sessions, automatically punctuated (Sections S3.03 and S4). The neuroprosthesis could also send the sentence to the participant's personal computer by acting as a Bluetooth keyboard, which allowed him to use it for activities such as writing email (Section S1.08). Sampled phonemes and words used for decoder training were accumulated over the course of the study (Fig. S4).

### Decoding Speech

No microphone input was used for decoding, and we found no evidence of acoustic or vibration-related contamination in the recorded neural signals (Section S4.03 and Fig. S7). Every 80 msec, the activity from the cortical recordings was used to predict the most likely English phoneme being attempted (Section S2 and Figs. S5 and S6). Phoneme sequences were then combined into words with the use of an openly available language model.[21] Next, we applied two additional open-source language models to translate the sequence of words that had initially been predicted from the neural activity into the most likely English sentence (Section S3 and Fig. S8), as described in a previous report.[21] Data from multiple days were combined to continuously calibrate the decoder (Section S2 and Figs. S9 and S10).

### Evaluation

We used two measures to analyze the speech-decoding performance that were consistent with those used in previous studies of speech decoding[17,19,21]: phoneme error rate and word error rate. These measures are calculated as the ratio of the number of phoneme or word errors to the total number of phonemes or words expected to be decoded, with the results expressed as percentages. An error is defined as the need for an insertion, deletion, or substitution in order to have the decoded sentence match the intended sentence (e.g., the prompted text in the copy task).

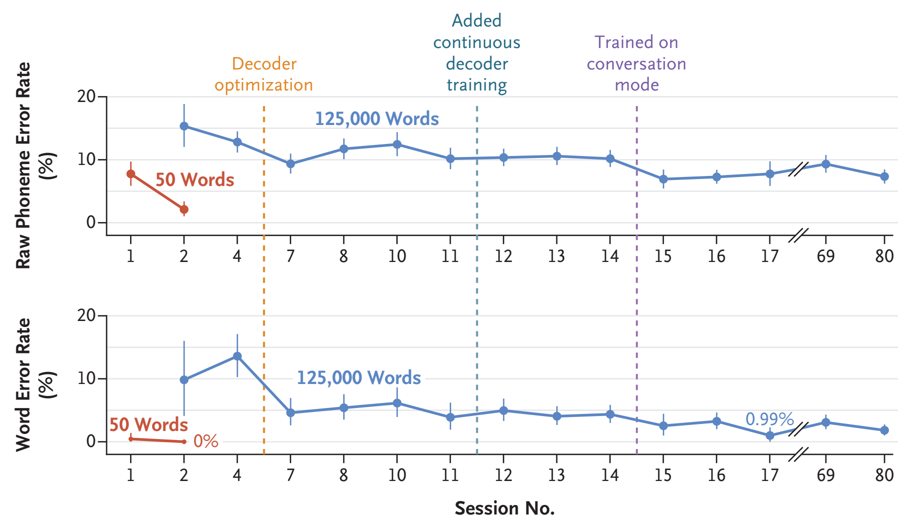
*Videos showing use of the neuroprosthesis are available at NEJM.org*

The phoneme error rate can be understood as the ability of the system to translate cortical neural activity into phonemes without language models ("raw" phoneme error rate in the figures), and the word error rate can be understood as an estimate of overall communication accuracy.

When accuracy was evaluated during the copy task, the correct text was the prompted text shown to the participant. When accuracy was evaluated during participant-initiated conversation, we used a combination of methods to identify the intended sentence, including asking the participant after the session what he had meant to say (Section S1.09). We also evaluated estimated sentence-level accuracy by having the participant use an eye tracker to select on-screen buttons corresponding to whether the preceding output text was "100% correct," "mostly correct," or "incorrect" (Section S1.08). The blocks during which data were collected were either "training blocks," in which data were collected for decoder training and optimization, or predetermined "evaluation blocks," which were used to measure and report performance. Error rates were aggregated over all evaluation sentences for each session (Section S1.09). The first-ever closed-loop block (during session 1) was excluded from evaluation because the experience of using the system elicited tears of joy from the participant and his family as the words he was trying to say appeared correctly on screen. (We confirmed that this affective display was concordant with his emotional state and not a pseudobulbar phenomenon.)

### STATISTICAL ANALYSIS

Results for each analysis are reported as means and either 95% confidence intervals or standard deviations. Confidence intervals were estimated by randomly resampling each data set 10,000 times with replacement and have not been adjusted for multiplicity. The measures used for the evaluation of decoding performance — the phoneme error rate and word error rate, both of




| Session No. | 1 | 2 | 4 | 7 | 8 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 69 | 80 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Days since Implantation | 25 | 27 | 34 | 46 | 48 | 69 | 74 | 76 | 81 | 83 | 88 | 90 | 95 | 223 | 244 |
| Cumulative Hours of Data | 0.5 | 1.9 | 4.0 | 6.6 | 8.0 | 8.9 | 10.1 | 11.5 | 12.2 | 13.4 | 15.9 | 17.6 | 18.8 | 80.9 | 94.5 |

**Figure 2. Online Speech-Decoding Performance.**

Phoneme error rates (top) and word error rates (bottom) are shown for each session for two vocabulary sizes (50 and 125,000 words). These measures are calculated as the ratio of the number of phoneme or word errors to the total number of phonemes or words expected to be decoded, with the results expressed as percentages. The mean aggregate error rate across all evaluation sentences is shown for each session; error bars indicate the 95% confidence intervals. The hours row beneath the horizontal axis indicates the cumulative hours of neural data used to train the speech decoder for that session. Vertical dashed lines indicate times when decoder improvements were introduced. Fig. S20 shows phoneme and word error rates for individual blocks.

which were evaluated with the use of the Levenshtein distance (an estimate of the minimum number of edits required to correct a sequence) — were chosen before the start of data collection (Section S1.09).

## RESULTS

### ONLINE SPEECH-DECODING PERFORMANCE

In the first session, the participant attempted to speak prompted sentences constructed from a 50-word vocabulary.[17] We recorded 213 sentences over 30 minutes of the copy task; these sentences were used to calibrate the speech neuroprosthesis. Next, we decoded the participant's cortical activity in real time as he tried to speak. The neuroprosthesis decoded his attempted speech with a word error rate of 0.44% (95% confidence interval [CI], 0.0 to 1.4), or 99.6% accuracy. We retested decoding with the 50-word vocabulary in the second research session, in which all the participant's attempted sentences were decoded correctly (word error rate of 0%) (Fig. 2).

In this second research session, we expanded the vocabulary of the neuroprosthesis from 50 words to more than 125,000 words, which encompasses the majority of the English language.[26] We collected an additional 260 sentences of training data over 1.4 hours. After training on these additional sentences, the neuroprosthesis decoded the participant's attempted speech with a copy-task word error rate of 9.8% (95% CI, 4.1 to 16.0), or 90.2% accuracy (Fig. 2). Performance continued to improve in subsequent research sessions as we collected more training data and adapted innovations for incorporating new data more effectively[27] (Sections S2 and S3). The neuroprosthesis achieved a word error rate of 2.5% (95% CI, 1.0 to 4.5) by session 15, and accuracy was maintained at or near 97.5% through session 84, more than 8 months after implantation. The mean decoding performance for the copy task in the final five evaluation sessions was a word error rate of 2.5% (95% CI, 2.0 to 3.1) at the participant's self-paced speaking rate of 31.6 words per minute (95% CI, 31.2 to 32.0) (Fig. S1B), with individual daily mean word error rates ranging from 1.0 to 3.3% (Table S3). The communication rate of the neuroprosthesis exceeded the participant's standard means of communication using a head mouse or skilled interpreter (Fig. S1A).

The performance of neural decoding was maintained across days (Figs. S14 and S15). The arrays in the ventral premotor cortex and middle precentral gyrus contributed most to decoding accuracy (Fig. S16). The neuroprosthesis decoded words it was not explicitly trained with (Fig. S18) and worked across different attempted speaking amplitudes, including nonvocalized speech (Fig. S19).

### CONVERSATIONAL SPEECH WITH THE USE OF THE BRAIN—COMPUTER INTERFACE

During conversational speech, the neuroprosthesis detected when the participant started or stopped speaking (Section S1.08 and Fig. S21). When evaluated offline with copy-task sentences (exercises in which we knew when he was or was not trying to speak), the system falsely detected that he wanted to speak for less than 1% of sentences (Fig. S22). In addition, the participant had the option to use an eye tracker for selecting actions (Fig. 3) to finalize and read a sentence aloud, to indicate whether the neuroprosthesis output was correct, or to initiate a mode in which he could spell out words letter-by-letter by attempting to say those letters. This option was useful for situations in which words were not correctly predicted by the decoder — for example, because they were words that were not in the vocabulary, such as certain proper nouns.

The participant's first use of the neuroprosthesis for conversational communication with his family is shown in Video 3 (additional transcripts are provided in Fig. S23 and Table S4). In subsequent sessions, he used the neuroprosthesis for personal conversations (e.g., Videos 4 and 5), communicating a total of 22,679 sentences during 72 sessions (out of 84 total sessions) over 8.4 months (248.3 cumulative hours) (Fig. 4A). The mean word error rate during selected sessions in conversation mode was 3.7% (95% CI, 3.3 to 4.3) (Fig. 4B), and the participant reported that 52.9% and 32.3% of all sentences in conversation mode were decoded correctly and mostly correctly, respectively (Fig. 4C). The longest continuous period during which the speech neuroprosthesis was used in conversation mode was 7.7 hours. Across the 29 session days during which the participant used the neuroprosthesis solely for personal communication, he requested that the decoder be recalibrated and updated on three occasions because of a larger number of errors than he was accustomed to. Each calibration took approximately 7.5 minutes, during which 20 copy-task sentences were displayed to provide training labels and thereby rapidly update the decoder.

The participant used conversation mode to perform activities ranging from talking to the research team, family, and friends to performing his occupation by participating in videoconferencing meetings and writing documents and email messages. Using the neuroprosthesis, the participant shared his perspectives on using the system (Table S4).

## DISCUSSION

Beginning on the first day of device use, 25 days after implantation, a brain-to-text speech neuroprosthesis with 256 cortical recording sites in the left precentral gyrus accurately decoded intended speech in a man with severe dysarthria due to ALS. He communicated using a 125,000-word vocabulary on the second day of use. Within 16 cumulative hours of use, the neuroprosthesis incorrectly identified only 2.5% of attempted words. To contextualize this error rate, the state of the art for English automated speech recognition (e.g., smartphone dictation) has an approximate 5% word error rate[28] and able-bodied speakers have a 1 to 2% word error rate[29] when reading a paragraph aloud.

The study participant used the speech neuroprosthesis to converse with family, friends, health care professionals, and colleagues at a rate of approximately 32 words per minute. His regular means of communication without a neuroprosthesis involved either having expert caregivers interpret his severely dysarthric speech or the use of a head mouse with point-and-click selections on a computer screen. The investiga-

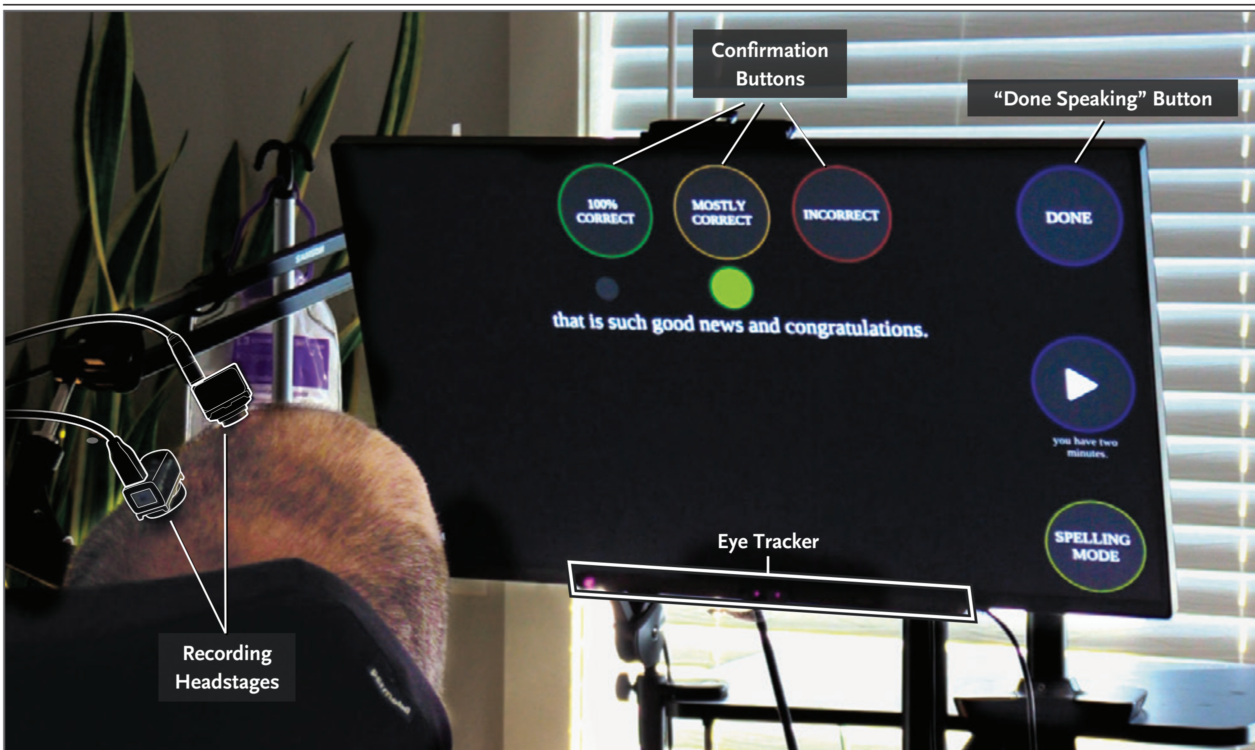



**Figure 3. Conversation-Mode User Interface.**

A photograph of the participant using the speech neuroprosthesis in conversation mode is shown. The neuroprosthesis detected when the participant was trying to speak solely on the basis of neural activity and concluded either after 6 seconds of speech inactivity or on his optional activation of an on-screen button through eye tracking. After the decoded sentence was finalized, the participant selected on-screen confirmation buttons through eye tracking to indicate whether the decoded sentence was correct.

tional system became his preferred way to communicate with our research team, and he used it on his own time; however, a researcher's assistance was required to connect and launch the system. The participant and his family indicated that the voice of the system resembled his own.

This study showed a reduction in the quantity of training data required to achieve high-accuracy decoding as compared with our previous study,[21] in which performance was tested starting 113 days after implantation and 16.8 hours of training data collected over 15 days was used to achieve a word error rate of 23.8%. Another previous speech neuroprosthesis required 17.7 hours of training data collected over 13 days to reach a word error rate of 25.5%.[19]

In addition to recording from two arrays in the putative ventral portion of the speech premotor cortex, as in a previous study,[21] we also implanted one array each in two areas for which previous reports of recordings with multielectrode arrays are lacking: Brodmann area 4 (primary motor cortex, which in humans is often in the central sulcus[23] and thus largely not accessible with microelectrode arrays) and area 55b.

This study involved a single participant, and whether similar results can be expected in future users is uncertain. Moreover, this patient had some residual audible speech function that may have contributed to the results in this study. The durability of the system as ALS progresses has not been extensively studied, and we cannot comment on the use of this system in patients with other disorders.

In a participant with ALS, a rapidly usable and accurate restoration of speech-based communication with an extensive vocabulary was enabled by an intracortical neuroprosthesis.

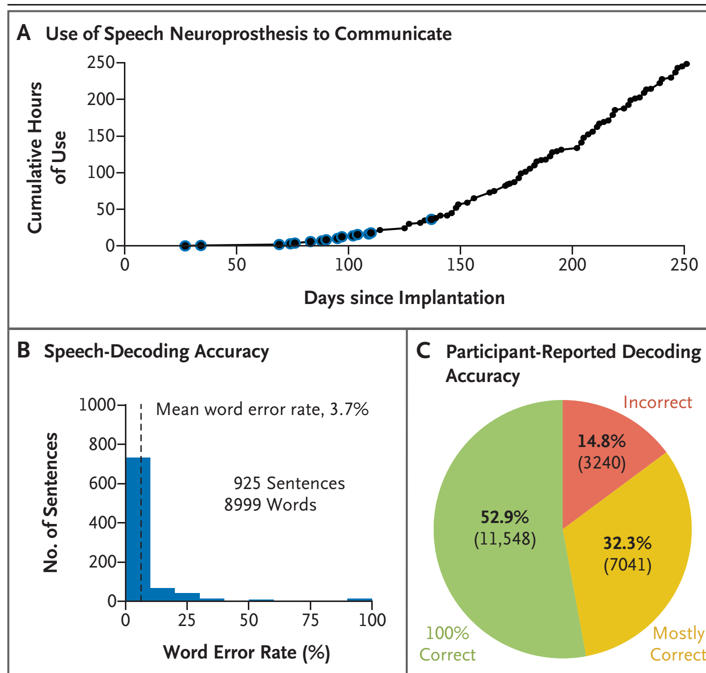



**Figure 4. Use of the Neuroprosthesis for Participant-Initiated Speech.**

Panel A shows the cumulative hours that the participant used the speech neuroprosthesis to communicate during structured research sessions and personal use. For the sessions outlined in blue, conversation-mode decoding accuracy was quantified as shown in Panel B. Panel B shows a histogram of speech decoding accuracy in conversations for the 925 sentences with known true labels (Section S1.09). The mean word error rate was 3.7% (95% CI, 3.3 to 4.3). Panel C shows participant-reported decoding accuracy for each sentence (which is different from the word error rate) across all conversation-mode data (21,829 sentences).

Supported by an ALS Pilot Clinical Trial Award (AL220043) from the Office of the Assistant Secretary of Defense for Health Affairs, a New Innovator Award (NIH 1DP2DC021055) from the National Institutes of Health and managed by the National Institute on Deafness and Other Communication Disorders, a postdoctoral fellowship funded by the A.P. Giannini Foundation, a grant (872146SPI) from the Simons Collaboration for the Global Brain, the Searle Scholar Program, the Burroughs Wellcome Fund, the University of California, Davis, a grant (NIH U01DC017844) from the National Institute on Deafness and Other Communication Disorders, a grant (A2295-R) from the Department of Veterans Affairs Rehabilitation Research and Development Service, the Howard Hughes Medical Institute, and the Wu Tsai Neurosciences Institute at Stanford University.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

**REFERENCES**

1. Papathanasiou I, Coppens P. Aphasia and related neurogenic communication disorders. Burlington, MA: Jones and Bartlett Learning, 2017.

2. Katz RT, Haig AJ, Clark BB, DiPaola RJ. Long-term survival, prognosis, and life-care planning for 29 patients with chronic locked-in syndrome. Arch Phys Med Rehabil 1992;73:403-8.

3. Lulé D, Zickler C, Häcker S, et al. Life can be worth living in locked-in syndrome. Prog Brain Res 2009;177:339-51.

4. Bach JR. Amyotrophic lateral sclerosis: communication status and survival with ventilatory support. Am J Phys Med Rehabil 1993;72:343-9.

5. Koch Fager S, Fried-Oken M, Jakobs T, Beukelman DR. New and emerging access technologies for adults with complex communication needs and severe motor impairments: state of the science. Augment Altern Commun 2019;35:13-25.

6. Luo S, Rabbani Q, Crone NE. Brain-computer interface: applications to speech decoding and synthesis to augment communication. Neurotherapeutics 2022;19:263-73.

7. Herff C, Heger D, de Pesters A, et al. Brain-to-text: decoding spoken phrases from phone representations in the brain. Front Neurosci 2015;9:217.

8. Kellis S, Miller K, Thomson K, Brown R, House P, Greger B. Decoding spoken words using local field potentials recorded from the cortical surface. J Neural Eng 2010;7:056007.

9. Mugler EM, Patton JL, Flint RD, et al. Direct classification of all American English phonemes using signals from functional speech motor cortex. J Neural Eng 2014;11:035015.

10. Ramsey NF, Salari E, Aarnoutse EJ, Vansteensel MJ, Bleichner MG, Freudenburg ZV. Decoding spoken phonemes from sensorimotor cortex with high-density ECoG grids. Neuroimage 2018;180:301-11.

11. Anumanchipalli GK, Chartier J, Chang EF. Speech synthesis from neural decoding of spoken sentences. Nature 2019;568:493-8.

12. Moses DA, Leonard MK, Makin JG, Chang EF. Real-time decoding of question-and-answer speech dialogue using human cortical activity. Nat Commun 2019;10:3096.

13. Stavisky SD, Willett FR, Wilson GH, et al. Neural ensemble dynamics in dorsal motor cortex during speech in people with paralysis. Elife 2019;8:e46015.

14. Stavisky SD, Willett FR, Avansino DT, Hochberg LR, Shenoy KV, Henderson JM. Speech-related dorsal motor cortex activity does not interfere with iBCI cursor control. J Neural Eng 2020;17:016049.

15. Berezutskaya J, Freudenburg ZV, Vansteensel MJ, Aarnoutse EJ, Ramsey NF, van Gerven MAJ. Direct speech reconstruction from sensorimotor brain activity with optimized deep learning models. J Neural Eng 2023;20:056010.

16. Guenther FH, Brumberg JS, Wright EJ, et al. A wireless brain-machine interface for real-time speech synthesis. PLoS One 2009;4(12):e8218.

17. Moses DA, Metzger SL, Liu JR, et al. Neuroprosthesis for decoding speech in a paralyzed person with anarthria. N Engl J Med 2021;385:217-27.

18. Metzger SL, Liu JR, Moses DA, et al. Generalizable spelling using a speech neuroprosthesis in an individual with severe limb and vocal paralysis. Nat Commun 2022;13:6510.

19. Metzger SL, Littlejohn KT, Silva AB, et al. A high-performance neuroprosthesis for speech decoding and avatar control. Nature 2023;620:1037-46.

20. Luo S, Angrick M, Coogan C, et al. Stable decoding from a speech BCI enables control for an individual with ALS without recalibration for 3 months. Adv Sci (Weinh) 2023;10(35):e2304853.

21. Willett FR, Kunz EM, Fan C, et al. A high-performance speech neuroprosthesis. Nature 2023;620:1031-6.

22. Yuan J, Liberman M, Cieri C. Towards an integrated understanding of speaking rate in conversation. In: Interspeech 2006, September 17–21, 2006. Pittsburgh: International Speech Communication Association, 2006:1795-Mon3A3O.1 (https://doi.org/10.21437/Interspeech.2006-204).

23. Glasser MF, Coalson TS, Robinson EC, et al. A multi-modal parcellation of human cerebral cortex. Nature 2016;536:171-8.

24. Silva AB, Liu JR, Zhao L, Levy DF, Scott TL, Chang EF. A neurosurgical functional dissection of the middle precentral gyrus during speech production. J Neurosci 2022;42:8416-26.

25. Ali YH, Bodkin K, Rigotti-Thompson M, et al. BRAND: a platform for closed-loop experiments with deep network models. J Neural Eng 2024;21:026046.

26. Godfrey JJ, Holliman EC, McDaniel J. SWITCHBOARD: telephone speech corpus for research and development. In: ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing, 23–26 March 1992. Vol. 1. San Francisco: Institute of Electrical and Electronics Engineers, 1992:517-520 (https://ieeexplore.ieee.org/document/225858).

27. Fan C, Hahn N, Kamdar F, et al. Plug-and-play stability for intracortical brain-computer interfaces: a one-year demonstration of seamless brain-to-text communication. Adv Neural Inf Process Syst 2023;36:42258-70.

28. Tüske Z, Saon G, Kingsbury B. On the limit of English conversational speech recognition. May 3, 2021 (https://doi.org/10.48550/arXiv.2105.00982). preprint.

29. Thomson DR, Besner D, Smilek D. In pursuit of off-task thought: mind wandering-performance trade-offs while reading aloud and color naming. Front Psychol 2013;4:360.