Speaking requires coordinating articulator muscles with exquisite timing and precision. Understanding how the sensorimotor system accomplishes this requires studying its neural underpinnings. Neural measurements are also critical for identifying the causes of speech disorders and building brain-computer interfaces (BCIs) to restore speech. Speech is a uniquely human behavior, which makes electrophysiological investigations challenging. Previous direct neural recordings during speech have come from electrocorticography (ECoG) or single-unit (SUA) recordings from penetrating electrodes during clinical epilepsy treatment. Such studies have begun to characterize motor cortical dynamics underlying speech, but not at the finer spatiotemporal scale uniquely afforded by high-density intracortical recordings often used in animal reaching studies.

We had the opportunity to study speech production at this detailed resolution by recording from multielectrode arrays previously chronically placed in human motor cortex as part of the BrainGate2 BCI clinical trial for people with paralysis. The spiking rate of neurons in dorsal ‘arm areas’, where speech-related activity has not previously been reported, modulated during speaking. This finding challenges whether the conventional model of a ‘motor homunculus’ somatotopy extends to the single-neuron scale. It is, however, consistent with known links between hand and speech networks that may reflect hand-mouth coordination and an evolutionary relationship between manual and articulatory gestures. Two neural population dynamics features previously reported for arm movements were also present during speaking: a large initial condition-invariant signal, followed by rotatory dynamics. This suggests that common neural dynamical motifs may underlie movement of arm and speech articulators. Lastly, spoken words and phonemes could be accurately decoded from single trials, demonstrating the potential utility of intracortical recordings for BCIs to restore speech.

**Finding 1: Dorsal motor cortex neurons respond during speaking and orofacial movements.** Participant ‘T5’ performed a cued speaking task in which he heard a phoneme or word cue played from a computer speaker and spoke back that sound after hearing a go cue (Fig. 1A). We recorded SUA action potentials as well as potentially multiunit threshold-crossing spikes (TCs) and LFPs from electrodes in dorsal motor cortex (1B). This same dorsal neural population was previously shown to modulate during attempted arm and hand movements. Here we found that these neurons’ firing rates strongly modulated during speech (1C). Across the arrays, 13/22 neurons and 73/104 electrodes’ TCs significantly responded during speaking at least one phoneme, with most responding to multiple phonemes (median = 4). This modulation was much greater when speaking than hearing sounds, and the same neural population was also active when T5 made orofacial movements of the mouth, lips, and tongue (1D). This suggests a broad encoding scheme related to motor cortical control of the speech articulators, consistent with previous ECoG findings in ventral motor cortex.
predicted individual trials' speaking reaction times ($r = -0.14$, $p<0.05$, linear regression).

Previous monkey\textsuperscript{4,9} and human\textsuperscript{5} studies found that subsequent peri-movement population activity during arm and hand movements is characterized by orderly rotatory dynamics. We saw similar neural state rotations during speaking (2C). These results suggest that motor cortex may ubiquitously deploy these dynamical motifs – a large transient input that “kicks” the network into a different state from which activity lawfully evolves according to rotatory dynamics – across a variety of different behaviors (e.g., speech and reaching) to construct the desired muscle activity from an oscillatory basis set.

An important open question for these findings is to what extent these results were influenced by potential cortical remapping due to tetraplegia. Definitively resolving this ambiguity would require intracortical recording from this eloquent brain area in able-bodied people.

Finding 3: Spoken words and phonemes can be decoded from intracortical activity.

Examining the single-trial neural population activity during speaking revealed strong clustering by phoneme, with phonemes further grouping by phonetic similarities (3A). With an eye to evaluating the possible utility of these neural signals for speech BCIs, we quantified how well they could be used to classify which sound was being spoken on a single-trial basis. Prediction accuracy amongst nine phonemes, plus silence, was 84.6% (3B; shuffled chance accuracy was 10.1%). Prediction between ten short words, plus silence, was 83.5% accurate (chance was 9.1%).

These results compare favorably to previously published decoding accuracies using ECoG\textsuperscript{1}, despite our presumably suboptimal dorsal recording locations. This suggests a promising path forward for speech BCIs using intracortical electrophysiology, which can take advantage of an anticipated dramatic scale-up in the number of neurons that new, higher-density sensors will be able to record.