Grasp, Speech, and Internal Speech Representation in the Human Cortical Grasp Circuit

Thesis by

Sarah Kim Wandelt

In Partial Fulfillment of the Requirements for the

Degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2023

(Defended April 10 2023)

Sarah Kim Wandelt

ORCID: 0000-0001-9551-8491

ACKNOWLEDGEMENTS

First, I would like to thank my advisor Richard Andersen for this exciting and unique research experience. What began as a six-month Master's project transformed into a five-year PhD journey, and I feel incredibly fortunate to have had the opportunity to learn and grow in this lab. Thank you for your mentorship over the years and for giving me the academic freedom to pursue new research directions. I would also like to thank my committee members Ralph Adolphs, Ueli Rutishauser, and Yisong Yue for providing me with insightful comments and feedback.

I was incredibly fortunate to have been mentored by two brilliant and kind scientists, Spencer Kellis and David Bjånes. Spencer, you provided me with a strong foundation to succeed in this field, always answering my abundance of questions and offering guidance and advice on my projects, not to mention the chocolate and donuts. You gave me the confidence to pursue a PhD and inspired me to apply to Caltech. David, thank you being such a great, smart, and fun person to work with. Thank you for all the helpful discussions we've had over the years, be it about science or life, and your excellent inputs on my projects. Thank you for helping me become a better writer, always championing my work, and believing in me even when I doubted myself. I would also like to thank the rest of the team: Isabelle Rosenthal for being an amazing fellow graduate student and friend, and Luke Bashford as an additional mentor for insightful discussions and advice over the years.

I would like to offer my heartfelt thanks to the research participants with whom I had the honor to work. Without their generous contributions of their time and effort, none of these projects would have been possible.

I would also like to express my gratitude to my PhD cohort in the Andersen Lab: Kelly Kadlec, Charles Guan, and Whitney Griggs. Working alongside such driven, helpful, and kind people has been a privilege, and has made my PhD experience here so enjoyable. I would also like to thank the rest of the lab family, who have supported me as friends or provided advice over the years: Sumner Norman, Srinivas Chivukula, Varun Wadia, HeyongChan Jo, Vasileios and Sofia Christopoulos, Tyson Aflalo, Jorge Gamez, Kelsie Pejsa, and Viktor Shcherbatyuk.

I've met some truly amazing people during my time at Caltech, be it through volleyball, the Neurotechers, or other social activities. I am deeply grateful to each of them for making my Caltech experience so special, through shared experiences such as social outings, barbeques, hikes, beach trips, and more. Katrina, you were my first close friend at Caltech and we have had so many fun adventures together since then. I am excited for many more to come! Zsofi, you've been such a caring friend and source of support. I am thankful for all the times you've been here for me.

I owe a great deal of thanks to my family. Maman et papa, je tiens à vous exprimer toute ma gratitude pour l'aide précieuse que vous m'avez apportée au fil des ans, en me guidant depuis mon enfance jusqu'à l'obtention de ce doctorat. Votre exemple en mettant en avant l'importance de l'éducation et de la connaissance a été un facteur clé de ma réussite, et vous m'avez toujours soutenu dans mes projets. Grâce à votre soutien, j'ai eu les opportunités qui m'ont permis de décrocher ce doctorat, sans lesquelles je n'aurais pas pu y arriver. Leila und Ines, ihr habt mich immer in meinem Leben unterstützt und ich bin so dankbar, euch als Schwestern zu haben. Ines, du hast mich mein ganzes Leben begleitet, und dafür werde ich dir nie genug danken können. Ich liebe euch alle so sehr.

I have been fortunate to meet my partner, Konrad Pilch, here at Caltech. I have never met anyone as kind and special as you, with your unparalleled ability to make me smile. You've been here for me through sickness and bad times, and I know that I can always count on you. I could not have achieved my PhD without you.

ABSTRACT

The ability to move freely and to connect with others through communication is invaluable for human independence. In this thesis, we explore how brain-machine interfaces (BMIs) can help patients affected by movement or speech deficits to recover lost human experiences. This work builds on previous findings indicating that premotor and posterior parietal areas are involved in movement generation and language processes. These higher-level brain areas do not only engage in movement execution, but also during planning, representing rich behavioral patterns that can be leveraged for BMI applications. In this work, we investigated how the ventral premotor cortex (PMv), the posterior parietal cortex (PPC), and the sensorimotor cortex (S1) represent grasp and speech processes at a single-neuron level. Using multielectrode Utah arrays, neuronal populations were recorded in tetraplegic human participants. We found that the supramarginal gyrus (SMG), PMv and S1 significantly encode motor imagery of grasping. By studying the cognitive processes underlying neural activity during the cue phase of grasping, we found a transition from cue-modality dependent to cue-modality independent grasp representation in SMG, the anterior intraparietal cortex (AIP) and PMv. Our findings suggest SMG integrates audio, written, and image cue modalities, but more similarly represents audio and written cues, indicating language involvement. We confirmed this hypothesis by demonstrating that SMG encodes spoken words, engaging different motor plans for speech compared to grasping even when the semantic content remained unchanged. These results suggest a BMI could be trained to decode both grasp motor imagery and speech from one brain area. Lastly, we showed that SMG is highly involved in language processes, modulating for written word recognition, auditory tones, vocalized speech, and internal speech. As a proof-of-concept, we built a real-time internal speech BMI from signals recorded in SMG that can decode eight words with high accuracy. This work is the first of its kind, demonstrating internal speech can be robustly decoded from an implant in a single brain area. We find high neural SMG generalization between seeing a written word, saying it internally and vocalizing the word, suggesting shared cognitive functions between different language processes. Furthermore, words in different languages are represented in SMG. This thesis advances the BMI field by providing a better understanding of the neural processes that underly grasp motor imagery and language. To summarize, our findings suggest that studying higher-level brain areas can lead to the development of more effective and versatile brain-machine interfaces.

PUBLISHED CONTENT AND CONTRIBUTIONS

"Decoding grasp and speech signals from the cortical grasp circuit in a tetraplegic human" **Sarah K. Wandelt**, Spencer Kellis, David. A. Bjånes, Kelsie Pejsa, Brian Lee, Charles Liu, Richard A. Andersen, Neuron 110 (11), 1777-1787. e3 https://doi.org/10.1016/j.neuron.2022.03.009

- Contributions: study design, developmental of experimental tasks, data acquisition, methodology, formal analysis, interpretation of results, writing
- Included in Chapters 3 and 5 of this thesis.

"Online internal speech decoding from single neurons in a human participant" **Sarah K. Wandelt**, David A. Bjånes, Kelsie Pejsa, Brian Lee, Charles Liu, Richard A. Andersen, medRxiv 2022.11.02.22281775; doi: https://doi.org/10.1101/2022.11.02.22281775

- Contributions: study design, developmental of experimental tasks, data acquisition, methodology, formal analysis, interpretation of results, writing
- Included in Chapter 6 of this thesis.

Table of Contents

LIST OF ILLUSTRATIONS AND/OR TABLES

NOMENCLATURE

SMG	Supramarginal gyrus
PMv	Ventral premotor cortex
S1	Primary somatosensory cortex
PPC	Posterior parietal cortex
M1	Primary motor cortex
AIP	Anterior intraparietal cortex
IPS	Intraparietal Sulcus
BMI	Brain-Machine Interface
ECoG	Electrocorticography
EMG	Electromyography
fMRI	Functional magnetic resonance imaging
ROI	Region of interest
sEEG	Stereo-electroencephalography
S2	Participant in the study with implants in SMG, PMv, and S1
N1	Participant in the study with implant in AIP
SCI	Spinal cord injury
LFP	Local field potential
Spike	Action potential
Unit	A neuron
Spike sort	Sort action potentials through clustering methods into single neurons
PCA/ dPCA	Principal component analysis / demixed PCA
c.i.	Confidence interval

Chapter 1:
Introduction

Catching up with your family and friends while grabbing a steaming cup of coffee are ordinary activities of everyday life. For individuals affected by spinal cord injuries, or who suffered strokes or neurological diseases, these essential abilities can unexpectedly be stripped away. Those affected by movement disorders may find themselves unable to perform basic tasks due to paralysis or confinement to a bed. In severe cases, patients may even lose the ability to speak, rendering them isolated and confined to their own thoughts. Imagine, now, a technology that would allow the recovery of these crucial human experiences by controlling a robotic arm or communicating with just your mind. This is the hope of brain-machine interfaces.

In this thesis, I will explore the potential of brain-machine interfaces (BMIs) to restore critical human experiences to those who have lost them. Specifically, I will focus on BMI applications in the human cortical grasp circuit and investigate how this network of brain regions represents motor imagery of grasp and speech processes at a single neuron level. By studying lesser understood higher-level brain areas such as the posterior parietal cortex (PPC) and the ventral premotor cortex (PMv), I aim to understand the cognitive signals that can be recovered from these brain regions and offer new insights into the possible applications of intracortical BMIs.

In **Chapter 2** I will provide background information on invasive and non-invasive BMIs, and dive into the different components of the human cortical grasp circuit and their role in hand function and motor planning. I will summarize the role of the premotor and posterior parietal cortex as target sites for grasp BMI applications and recent efforts of decoding speech and internal speech from these areas.

In **Chapter 3**, we will show how the supramarginal gyrus (SMG), PMv, and the somatosensory cortex (S1) encode planning and execution of grasp motor imagery at a single neural level. We find all three regions were significantly modulated in response to motor imagery of five grasps. Additionally, the higher-level brain areas SMG and PMv also represent grasp-related activity when cued with a visual image, aligning with findings from nonhuman primate studies. This early cue phase activity is not

merely a visual response but also contains motor or semantic characteristics. Moreover, we demonstrated through a neuron-dropping analysis that neurons in SMG and PMv are similarly informative for grasp decoding, suggesting that excellent grasp classification accuracies can be obtained when recording from a substantial number of neurons in either area simultaneously. To summarize, this study highlights the potential of SMG and PMv for future grasp BMI applications.

Chapter 4 delves deeper into the cognitive processes represented during the cueing of grasps in higher-level brain areas of the cortical grasp circuit. Specifically, we investigated how sensory cue modalities affect motor imagery of five grasps, posing the following question: is grasp motor imagery invariable to the sensory cue modality? We cued tetraplegic patients with a visual image, an auditory cue, and a written cue in the SMG, PMv, and anterior intraparietal cortex (AIP). Our findings revealed a transition from modality-dependent representation during the cue phase to modality-invariant representation during motor imagery. Brain areas PMv and AIP exhibited cue-dependent neural modulation to the image cue, but not auditory or written cues, while SMG responded distinctively to all three cue modalities. Moreover, our analysis indicated that in SMG, auditory and written processes are more similarly represented than image processes, suggesting that SMG also plays a role in language processes.

In **Chapter 5**, we investigated SMG's, PMv's and S1's involvement in language by performing a vocalized speech task. Additionally, we compare these brain regions' activity to a grasp motor imagery task. Our findings show that SMG not only encodes grasp names, but also color names, indicating that the modulation of SMG to speech is not solely driven by action words. Furthermore, grasp and speech motor processes appeared to be distinctly represented in SMG, even when the semantic content remained the same. In the context of BMIs, these findings suggest we can train a decoder to represent both speech and grasp processes at the same time.

In **Chapter 6** we provided a proof-of-concept for a real-time internal speech BMI from signals recorded in an area of the SMG, which can decode eight different words with high accuracy. Indeed, an ideal speech BMI should represent movement-independent internal speech (also known as imagined or covert speech), which would allow locked-in patients to use the device. This work is the first of its kind, demonstrating internal speech can be robustly decoded from an implant in a single brain area. We find high neural SMG generalization between seeing a written word, saying it

internally, and vocalizing the word, suggesting shared cognitive functions between different language processes. We investigate phonetic and semantic representation by decoding words with identical semantic meanings as well as homophones and found evidence for both processes being represented within SMG.

This thesis advances the BMI field by increasing our understanding of the neural processes underlying grasp motor imagery and language. We show SMG not only robustly represents motor imagery but is also strongly involved in vocalized and internal speech processes. By building the first real-time internal speech BMI, we provide evidence that studying higher-level brain areas has important implications for the development of more effective and versatile brain-machine interfaces that can decode both grasp-related and language-related signals from the same brain area.

Chapter 2:
Background

BMI methods and techniques

Brain-machine interfaces (BMIs) are devices that interface with the central or peripheral nervous system and read out biological signals related to cognitive functions, bypassing the usual output modalities such as hand movement or speech. Typically, these signals are then used to control an external device, for instance by controlling a cursor on a screen, moving a robotic hand or displaying decoded words on a screen (Simeral et al. 2011; Hochberg et al. 2006; Collinger et al. 2013; Klaes et al. 2015; Moses et al. 2021). BMIs can also write information into the brain, allowing to evoke sensation through intracortical microstimulation (ICMS) (Flesher et al. 2016; Armenta Salas et al. 2018). For individuals who have experienced interruptions in their communication pathways from brain to limbs due to spinal cord injury, stroke, or disease, BMIs offers the hope of recovering some of their lost motor and sensory abilities.

One target group for BMI applications is people suffering from tetraplegia. Tetraplegia is defined by the loss of motor function and/or sensory function in all four limbs and the torso and can be caused by spinal cord injuries (SCI), stroke, or disease. Approximately 300,000 people in the US alone are thought to live with SCI injury, with about 18,000 new cases per year. Worldwide, this number is estimated to be close to 1 million (Ding et al. 2022). Close to 60% of SCI injuries lead to complete or incomplete tetraplegia, and less than 1% of patients will experience complete neurological recovery by the time of hospital discharge (National Spinal Cord Injury Statistical Center 2022). Another cause for severe loss of mobility is degenerative neurological diseases like amyotrophic lateral sclerosis (ALS) and brainstem strokes. In ALS, the progressive degeneration of motor neurons can even lead to a locked-in state (LIS), a condition that severely limits limb, speech, and facial movements. These patients are awake and conscious, unable to communicate or move, but may retain vertical eye movement (Smith and Delargy 2005; Norris et al. 1993; Chaudhary, Birbaumer, and Curado 2015). In 2015, around 16,500 people were reported to live with ALS in the US (Mehta 2018). For patients

affected by these conditions, tapping directly into the brain to read out their intentions could facilitate movement and communication.

BMIs offer a promising solution by capturing the electrical signals, also called action potentials or "spikes", that neurons use to communicate. These signals allow the processing and transmission of information throughout the brain and underlie our ability to perceive, think, learn, and execute complex behaviors. Various monitoring techniques exist to either directly or indirectly identify neural activity in the brain, and the acquisition methods can result in varying temporal and spatial resolution as well as quality of the obtained signal, which should be considered for real-time BMI applications.

Functional magnetic imaging (fMRI) indirectly measures neural activity by tracking changes in blood oxygenation levels in the brain due to task demands. While being non-invasive, fMRI has slow temporal and limited spatial resolutions, and is expensive to perform. Despite its limitation for real-time BMI applications, it can be useful for informing target brain areas for implantable BMIs. For instance, patients that meet the requirements to be scanned will often perform fMRI experiments to highlight which patch of the cortex is most active during tasks of interest (Tyson Aflalo et al. 2015; Armenta Salas et al. 2018).

Another non-invasive recording technique is electroencephalography (EEG). EEG records the voltage changes that are generated during neural communication through surface electrodes. While surgery is not required, the acquired signal is often subject to noise, has limited spatial resolution, and a lower signal-to-noise ratio than more invasive methods (Choi et al. 2018). As a result, these signals are insufficient to reconstruct natural movement and speech processes in patients affected by paralysis or strokes. For example, state-of-the art EEG-based imagined speech decoding performances in 2022 ranged from approximately 60 – 80% binary classification (Lopez-Bernal et al. 2022).

Magnetoencephalography (MEG) measures the electrochemical properties of neurons, providing similar temporal and spatial resolution as EEG (Malmivuo 2012). Similar to fMRI, they are currently expensive and immobile, limiting their use in BMI applications (Kauhanen et al. 2006; Dash, Ferrari,

and Wang 2020a). However, the neurotech company Kernel has been working on addressing these limitations by developing a portable MEG system based on optically pumped magnetometer (OPT) (Pratt et al. 2021), which might enable portable MEG-BMI systems in the future.

Invasive devices require the implantation of electrodes on the surface of the cortex through craniotomy. Electrocorticography (ECoG) places a grid of macro electrodes on the cortex, allowing coverage of a wide area and simultaneously offers increased signal quality and spatial resolution compared to EEG. Due to the size of the employed electrodes and their impedances, ECoG picks up local fluctuations of field potentials from multiple neural sources (Shokoueinejad et al. 2019). These local field potentials (LFP) contain useful information for movement and speech processes, as evidenced by successful ECoG-based BMIs (Moses et al. 2021; Anumanchipalli, Chartier, and Chang 2019; Kellis et al. 2010; Angrick et al. 2018; Hotson et al. 2016). However, they are not typically employed for the recording of action potentials of individual neurons in humans.

Stereotactic-EEG (sEEG) involves the implantation of penetrating depth electrodes using stereotactic guidance and allows sampling subcortical single-neuron activity. During sEEG, electrodes get inserted into targeted brain areas through small burr holes in the skull, which were shown to lead to less surgical complications than the craniotomies required for ECoG and multielectrode arrays (Iida and Otsubo 2017). While sEEG allows targeting deeper cortical regions, coverage per area is sparse, and is currently not in use for chronic recordings. Nonetheless, sEEG shows potential for BMI applications, and has been employed for grasp decoding and internal speech detection (Herff, Krusienski, and Kubben 2020; Meng Wang et al. 2020; Angrick et al. 2021).

Multielectrode arrays, such as the Utah array, involve placing densely spaced electrodes a few millimeters into the cortex. Each Utah array measures approximately 4 mm x 4 mm, allowing high-quality recordings of continuous electrical activity of the brain as well as the action potentials of individual neurons. With high signal-to-noise ratio, spatial, and temporal resolution, these arrays have shown great success in capturing information-rich signal variations required for high-quality decoding of movement related activity (Nicolas-Alonso and Gomez-Gil 2012; Brandman, Cash, and Hochberg 2017; Meijian Wang and Guo 2020). Invasive BMIs can be especially useful for patients that have suffered spinal cord injury (SCI) and are in the need of chronic BMI devices. For these

patients, cortical neural circuity remains intact, while the spinal cord lesion prevents communication between the brain and downstream muscles and nerves. Several groups, such as our own, have shown that useful neural signals can still be obtained seven years after microelectrode implantation without major side effects or complications (Mullin 2022, data from Chapter 6).

In this work, intracortical microelectrodes were used to obtain high quality recordings of movement and speech processes. After the neural signals are acquired, they undergo a series of processing steps, including filtering and amplification. Specifically, the signal obtained from multielectrode Utah arrays is filtered between 0.3-7500 Hz and thresholded to detect spiking activity. For BMI applications, the obtained signals are then mapped to behaviors using machine-learning algorithms. For example, these algorithms can learn to predict what hand gesture a participant is imagining. Once a model is trained, it can be used to control an external device, such as a robotic arm, based on the participant's intentions.

Figure 2-1A depicts a typical setup for a grasp or speech BMI experiments. A tetraplegic participant with implanted microelectrode arrays sits in front of a screen and is cued to perform an action such as motor imagery or speech vocalizations. The cue can be presented with different sensory modalities, including visual images of a grasp, auditory sounds, or written words. After a short delay, the participant performs the instructed action. Depending on the desired goal of the BMI application, the participant may imagine executing a specific grasp, attempt to reach an object, or vocalized a word. While the participant performs the task, neural responses are recorded and processed.

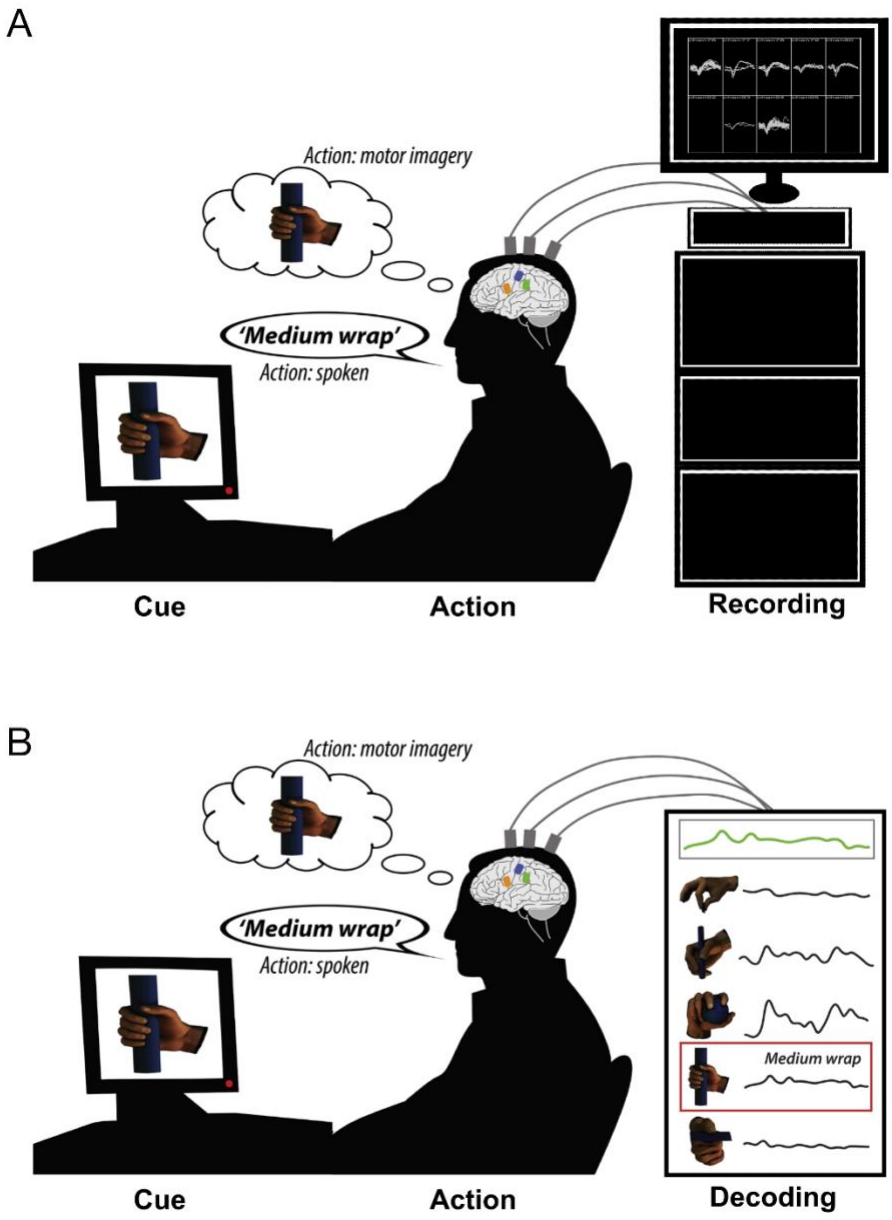

Figure 0-1 | Example setup of a BMI experiment. A) A tetraplegic participant implanted with microelectrode arrays is sitting in front of a screen and is cued with a visual image of an action to perform. After a short delay, the participant imagines perming the action, while neural responses are recorded and processed. **B)** During online control of then experiment, the action type is decoded from neural signals in real time.

Adapted from: Edmondson, Laura R., and Hannes P. Saal. "Getting a grasp on BMIs: Decoding prehension and speech signals." Neuron 110.11 (2022): 1743-1745.

Here, we can distinguish between offline and online BMIs. Offline, or so-called asynchronous BMIs, record neural activity while the participant performs the task but do not provide immediate control of an output device. The data is then analyzed post hoc to design machine-learning algorithms that

predict actions the participant may want to perform. Once a model is trained, it can be employed for online BMI applications (Figure 1B). In online BMI, neural data is received in real-time and translated into an action using the trained model. This approach allows for real-time control of a robotic arm or decoding of words for communication (Hochberg et al. 2012; Collinger et al. 2013; Klaes et al. 2015; Moses et al. 2021).

Beyond motor cortex: studying areas of the cortical grasp circuit for BMI applications

BMIs for grasp decoding

A crucial factor to consider when designing BMIs is the location of the implant, as it determines what type of cognitive processes can be recovered. People affected by tetraplegia often consider the restoration of hand function the most crucial skill they would like to regain (Anderson 2004). For grasp applications, implanting BMIs into brain regions involved in grasp planning and execution are therefore an obvious choice. The grasping and manipulations of objects is planned and executed through a complex neural system involving several brain regions that together are referred to as the cortical grasp circuit. Main areas involve parietal and premotor areas for grasp planning, the motor cortex for execution, and the sensorimotor cortex for sensory feedback. The first human target site for chronic motor BMI application was the primary motor cortex, and it has remained a popular target site ever since (Brandman, Cash, and Hochberg 2017; Hochberg et al. 2006; Collinger et al. 2013; F. R. Willett et al. 2021; Hochberg et al. 2012). Due to its proximity to output effectors, it contains a variety of useful signals for trajectory, reach, and grasp BMI applications (Brandman, Cash, and Hochberg 2017; Kalaska 2009; Rizzolatti, Luppino, and Matelli 1998). While the hand representation in the primary motor cortex (M1) is the final stage of grasp execution and generates the signals for the underlying finger joints for grasping, movement planning takes place in the higher-level premotor cortex and parietal areas. There, sensorimotor transformations for grasping are thought to occur (Andersen and Buneo 2002). Therefore, exploring brain regions outside the motor cortex may allow BMIs to represent richer behavioral patterns (Andersen, Aflalo, and Kellis 2019; Gallego, Makin, and McDougle 2022). In this work, we study three previously untested areas in human for grasp BMI applications: the ventral premotor cortex (PMv), the supramarginal gyrus (SMG), and the somatosensory cortex (S1).

In primate electrophysiology studies, both areas in the posterior parietal cortex (PPC) and the ventral premotor cortex (PMv, putative area F5 in primate) have shown selectivity for object features such as size, shape, and orientation (A. Murata et al. 2000; Akira Murata et al. 1997; Sakata 1995; Schaffelhofer, Agudelo-Toro, and Scherberger 2015a). Neurons recorded here show a variety of activation patterns. The neurons that process grasping intendent of visual input are called motor-dominant cells, while those that react stronger to grasping in well-lit conditions are called visuo-motor neurons. Additionally, visuo-dominant neurons process the visual cue even if no grasping is being performed (Sakata 1995; Taira 1998).

PPC is thought to play a central role in multisensory integration and to coordinate transformations at the earliest stages of movement planning. Indeed, parietal and premotor areas are not only active during grasp planning, but also at cue presentation (Townsend, Subasi, and Scherberger 2011; Klaes et al. 2015; Schaffelhofer and Scherberger 2016). Lesions in PPC lead to deficits in apraxia, an inability in planning movements, or optic ataxia, leading to inaccuracies in reaching to visual targets in the periphery. The presence of these goal-related signals for reaching could translate into faster decoding of movement intentions (Andersen et al. 2014a; Andersen and Buneo 2002; Bruni et al. 2017). Furthermore, a study found that many action variables are encoded in a partially mixed representation within PPC, allowing to decode movement intentions from most of the body in a small population of neurons (Zhang et al. 2017). PMv involvement in grasp planning and execution has been described on the single-neuron level in primates (A. Murata et al. 2000; Townsend, Subasi, and Scherberger 2011; Carpaneto et al. 2011; Schaffelhofer, Agudelo-Toro, and Scherberger 2015a; Schaffelhofer and Scherberger 2016; Bonini et al. 2010; Bruni et al. 2017), but it remains unclear whether these results translate to humans. These findings make parietal and premotor areas excellent alternative target sites for grasp BMI applications (Andersen, Aflalo, and Kellis 2019).

In human subjects, few papers have studied grasp and finger decoding in areas located outside the motor cortex. (Klaes et al. 2015) showed how single neurons in the anterior intraparietal cortex (AIP) encode five different grasp shapes in a participant with tetraplegia. Neurons were selectively active for specific hand shapes, but also strongly activated to the visual grasp cue. Offline, decoding of three different grasps achieved over 90% classification accuracy, while five different grasps could

be decoded with up to 65% accuracy. Online brain control of three grasps was also demonstrated, achieving above 80% classification accuracy. (Zhang et al. 2017) found neuronal modulation in AIP to hand squeezes. Recently, (Guan et al. 2022) achieved contralateral and ipsilateral finger movement decoding in two tetraplegic participants from AIP. Offline, right- and left-hand individual finger decoding accuracies of 70% and 66% were obtained, while online, six classes were decodable with 86% and 92%.

SMG is believed to play a crucial role in various grasp-related activities. It is hypothesized to have evolutionarily duplicated from human AIP specifically for tool use (Orban and Caruana 2014). A study conducted on human participants found that observation of tool use robustly activates left SMG. However, this response was not observed in non-human primates, even after training them how to use the involved tools (R. Peeters et al. 2009; R. R. Peeters, Rizzolatti, and Orban 2013). Furthermore, deficits in tasks involving tool usage are observed in patients with lesions near SMG (Goldenberg and Spatt 2009). Transcranial magnetic stimulation (TMS) over SMG while grasping objects for tool use leads to erroneous online corrections, further suggesting a causal role in tool use (McDowell et al. 2018). Studies have also confirmed SMG activity is modulated during grasping and manipulation of objects, reaching, and tool use (Sakata 1995; Filimon et al. 2009; Gallivan et al. 2013; Buchwald, Przybylski, and Króliczak 2018). (Johnson-Frey 2004a) also demonstrated SMG's involvement in both the planning and execution of pantomimed tool use. These characteristics underscore the rich potential of SMG as a source of grasp-related neural signals in the human cortex, that could be utilized for the BMI control of a prosthetic hand.

While the somatosensory cortex (S1) does not contribute to grasp planning per se, recent studies suggest it may serve as a promising target site for BMI applications in patients with tetraplegia. Human and NHP studies have shown that hand kinematics can be decoded during executed hand gestures (Branco et al. 2017) and before object grasping (Okorokova et al. 2020) respectively in S1. These findings suggest that neural signals may also be present during imagined movement (Zhang et al. 2017). Additionally, modulation of S1 neurons during motor imagery of reaching (Jafari et al. 2020) and decoding of imagined cutaneous and proprioceptive sensations (Bashford et al. 2021) have been demonstrated for the same participant whose data is the basis of this work. If grasp motor imagery can be reliably decoded, a single implant in S1 could facilitate a bidirectional BMI

capable of decoding grasp intentions and using ICMS to elicit somatosensations (Armenta Salas et al. 2018). Therefore, we evaluated S1 contributions to grasp motor imagery decoding.

BMIs for speech decoding

Patients with paralysis or condition like ALS may not only face challenges with movement, but also communication. In ALS, speech loss is often already apparent at time of diagnosis, and within 18 months, 60% of patient lose their ability to speak (Makkonen et al. 2018). This loss of speech is considered one of the most difficult aspects of the disease by ALS patients themselves (Hecht et al. 2002), underscoring the critical need for the development of high-performing assistive communicative devices. In recent years, the field of speech BMIs has been rapidly evolving, allowing to translate neural signals into texts (Moses et al. 2021; J. G. Makin, Moses, and Chang 2020) and acoustic speech features (Anumanchipalli, Chartier, and Chang 2019; Angrick et al. 2018). For review see (Luo, Rabbani, and Crone 2022; Cooney, Folli, and Coyle 2022).

Non-invasive recording techniques like fMRI, EEG, or magnetoencephalography (MEG) can be useful for identifying regions of the brain involved in vocalized and internal speech production (Dash, Ferrari, and Wang 2020b; Dash et al. 2020), however, they may lack the necessary spatiotemporal resolution, signal-to-noise ratio, and/or portability required to build an online speech BMI (Luo, Rabbani, and Crone 2022; Martin et al. 2018; Rabbani, Milsap, and Crone 2019).

When discussing speech BMI applications, it is important to distinguish the type of speech signals they are decoding (Figure 2-2). Indeed, while vocalized speech decoding has yielded excellent results, these findings have yet to translate to internal speech (see below). Therefore, it is still unclear if findings in vocalized speech decoding will translate to patients who have already lost part or all of their speech abilities.

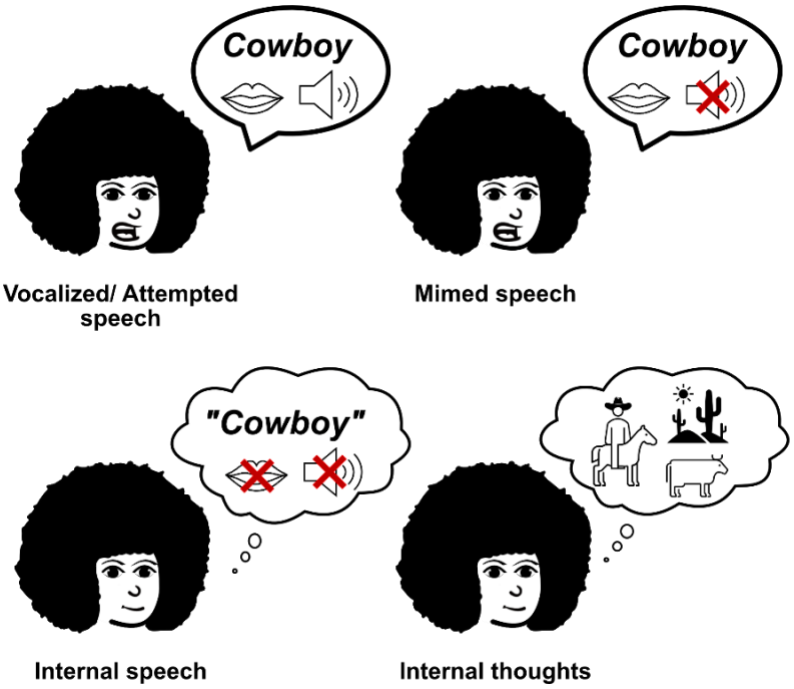

Figure 0-2 | Type of speech signals decoded by speech BMI applications.

Vocalized and attempted speech involve the production of sound through the vocal cords. Vocalized speech has successfully been decoded using a variety of intracortical BMIs, using ECoG (Kellis et al. 2010; J. G. Makin, Moses, and Chang 2020), sEEG (Angrick et al. 2018; Herff et al. 2019), and multielectrode arrays (Wilson et al. 2020; Stavisky et al. 2019). (Anumanchipalli, Chartier, and Chang 2019) demonstrated intelligible speech synthesis from neural ECoG features is possible by decoding spoken sentences from five participants who underwent intracranial monitoring for epilepsy. In a patient that had lost the ability to articulate speech but was still able to attempt speech, (Moses et al. 2021) decoded individual words from a vocabulary of 50 words. Using natural-language models and their vocabulary set, they were able to decode sentences with a median rate of 15.2 words per minute and an error rate of 25.6%.

Mimed speech, also sometimes referred to as silent speech, retains the same facial expressions and motor commands as vocalized speech, however no sound is being produced. Miming signals can also be used as an effective means of communication in speech BMIs (Bocquelet et al. 2016). In a recent study by (Metzger et al. 2022), it was shown that miming 26 words from the NATO phonetic alphabet enabled rapid and highly accurate spelling of sentences in the same participant as in the previous mentioned study.

Internal speech, also referred to as covert speech, imagined speech, or self-talk, refers to the process of silently talking to oneself in one's mind, without associated motor output. Importantly,

internal speech is distinct from internal thoughts that are more abstract and refer to a much broader range of mental processes occurring in one's mind. In contrast to vocalized and mimed speech, highly accurate real-time internal speech decoding has not yet been achieved. The main problem resides in the lack of external behavioral output. While vocalized and attempted speech produce acoustic signals that can be captured by a microphone, and mimed speech can be detected by using electromyography (EMG), no such signal exists for internal speech. However, these signals are required for the optimal training of speech decoding algorithm. These algorithms require knowing which timepoints in the neural signals correspond to spoken words or sentences. For example, the model employed by (Metzger et al. 2022) for real-time miming detection required supervised learning, where each datapoint in the neural data was labeled as either "rest", "speech preparation", "motor" or "speech". Furthermore, lower signal-to-noise ratio and decreases in cortical activation compared to vocalized speech make for particularly challenging internal speech decoding (Angrick et al. 2018; Martin et al. 2018; Luo, Rabbani, and Crone 2022; Proix et al. 2022).

Nonetheless, a few studies have shown above chance classification of internal speech in offline decoding analysis. (Pei, Barbour, et al. 2011) conducted a study on patients with ECoG grids implanted over frontal, parietal, and temporal regions, where participants silently read or vocalized written words displayed on a screen. The study successfully decoded vowels (37.5%) and consonants (36.3%) from internal speech, surpassing the chance level of 25%. In a similar study, (Ikeda et al. 2014) employed ECoG to decode three internally spoken vowels (chance level 33.3%) using frequencies in the beta band, with up to 55.6% accuracy from Broca area. Also using ECoG (Martin et al. 2018) investigated the decoding of six words during internal speech, achieving an average pair-wise classification of 58%, that reached up to 88% percent for the highest pair (chance level 50%). In a study conducted by (Dash, Ferrari, and Wang 2020b), the decoding of five internally spoken sentences using non-invasive MEG signals achieved up to 93% classification accuracy. However, the study had some limitations, including the considerable variation of the length of the tested sentences ("Goodbye" vs. "Do you understand me"), and the potential conflation of internal speech with speech preparation, both which may have contributed to higher classification accuracies. To our knowledge, the only study attempting to decode internal speech in real-time was recently conducted by (Angrick et al. 2021) using sEEG. While results were encouraging, as internal speech could be detected, the reconstructed audio was not yet discernable. Moreover, it is

important to note that the model utilized in the study required vocalized speech for training, which may not be feasible to obtain depending on the severity of the participant's condition.

Speech decoding in areas of the cortical grasp circuit.

Recent studies have shown that speech signals can be found in areas of the cortical grasp circuit. For instance, (Stavisky et al. 2019) and (Wilson et al. 2020) demonstrated speech decoding from the "hand knob" area in M1. Moreover, evidence for language processing has also been documented in PPC (Geranmayeh et al. 2012). (T. Aflalo et al. 2020) found a shared neural substrate for action verbs and observed actions, suggesting PPC may contribute to giving words their meaning. Additionally, (Zhang et al. 2017) found AIP modulation for spoken words. Investigating speech signals in the cortical grasp circuit could have promising applications, particularly for patients who suffer from both speech and movement deficits as in ALS. Implanting only one brain area that can represent both movement and speech processes would reduce surgery time and associated risks.

SMG involvement in speech processing

SMG involvement in language processing is apparent in several studies, however, to the best of our knowledge, its speech-related activity has not been characterized on the single-neuron level. Imaging studies have often observed preferred SMG activation for phonological over semantic decisions on visually presented words (C. J. Price et al. 1997; Seghier et al. 2004; Oberhuber et al. 2016). For instance, activation is stronger when participants have to decide on the number of syllables of a word, or if the presented word is an animal. Additionally, SMG activity also appears stronger for pseudowords (words composed of letters that conform to English orthographic pattern and are pronounceable, but have no semantic meaning), than for lexical words (Seghier et al. 2004). TMS over SMG can affect phonetic and semantic visual reading tasks (Stoeckel et al. 2009; Hartwigsen et al. 2010) and verbal working memory (Deschamps, Baum, and Gracco 2014). Furthermore, direct cortical electrostimulation of SMG was found to interfere with pseudoword reading (Roux et al. 2012). These findings provide evidence that SMG is involved in phonological aspects of language generation and could contain useful signals for speech BMIs.

SMG activation for speech production and during aspects of speech comprehension has been described by several studies (Cathy J. Price 2010). Using electrocorticography, (Pei, Leuthardt, et al. 2011a) found high-gamma SMG modulation during both vocalized and internal speech. Recently, (J.

G. Makin, Moses, and Chang 2020) decoded speech using ECoG signals, and found that electrodes located in SMG contributed to speech decoding. Furthermore, a study performed in people suffering from aphasia found lesions in SMG and its adjacent white matter affected inner speech rhyming tasks (Geva et al., 2011). These studies suggest that SMG is involved in various aspects of language processing, including word production, decoding, and phonological and semantic processing.

Overall, we covered how neurons in the ventral premotor cortex, the posterior parietal cortex, and the somatosensory cortex are implicated in grasp and speech generation. In this thesis, we will evaluate their potential for BMI applications by recording their activity with multielectrode Utah array and characterizing activity for grasp and speech processes at the single-neuron level.

Chapter 3:

Decoding grasp signals from the cortical grasp circuit in a tetraplegic human

The following chapter's contents are taken and adapted from Wandelt et al. 2022, with modifications done to fit the dissertation format. Work of this paper related to grasp processes is covered in this chapter. For work related to speech signals, see Chapter 5.

Wandelt, S. K., Kellis, S., Bjånes, D. A., Pejsa, K., Lee, B., Liu, C., & Andersen, R. A. (2022). Decoding grasp and speech signals from the cortical grasp circuit in a tetraplegic human. Neuron, 110(11), 1777-1787

Abstract

Tetraplegia from spinal cord injury leaves many patients paralyzed below the neck and unable to perform most activities of daily living. Brain-machine interfaces (BMIs) could give these patients greater independence by directly utilizing brain signals to control external devices such as robotic arms or hands. The cortical grasp network has been of particular interest because of its potential to facilitate the restoration of dexterous object manipulation. Neural activity from the cortical grasp network related to motor intentions for grasping was recorded in a tetraplegic patient in the supramarginal gyrus (SMG), the ventral premotor cortex (PMv), and the somatosensory cortex (S1). In high-level brain areas SMG and PMv, five imagined grasps were well represented by firing rates of neuronal populations during visual cue presentation. During motor imagery, grasps could be significantly decoded from all brain areas. At identical neuronal population sizes, SMG and PMv achieved similar highly significant decoding abilities, demonstrating their potential for grasp BMIs. These findings suggest that grasp signals can be robustly decoded at a single-unit level from the human cortical grasping circuit.

Introduction

The ability to grasp and manipulate everyday objects is a fundamental skill, required for most daily tasks of independent living. Functional loss of this ability, due to partial or complete paralysis from a spinal cord injury (SCI), can irrevocably degrade an individual's autonomy. People with tetraplegia have consistently rated recovery of hand and arm function as the highest priority for increasing their quality of life (Anderson 2004; Snoek et al. 2004).

Brain-machine interfaces (BMI) could give tetraplegic individuals greater independence by directly recording neural activity from the brain and decoding these signals to control external devices such as a robotic arm or hand (Tyson Aflalo et al. 2015). Intracortical BMIs use microelectrode arrays to capture the action potentials of individual neurons with a high signal-to-noise ratio (SNR) and high spatial resolution (Nicolas-Alonso and Gomez-Gil 2012). If placed in brain areas of the human grasp circuit, these devices are well suited to extract the neuronal signals supporting the control of a high-dimensional prosthetic hand (Collinger et al. 2013).

In this work, we evaluated the encoding of grasp motor imagery in human supramarginal gyrus (SMG), a sub region of the posterior parietal cortex (PPC), the ventral premotor cortex (PMv) and the primary sensory cortex (S1). These brain areas are key components of the cortical grasp circuit. PPC and PMv each encode complex cognitive processes, like goal and end-target directed signals (Tyson Aflalo et al. 2015), but, similar to M1, have also been shown to encode low level trajectory and joint-angle motor commands (Andersen et al. 2014; Schaffelhofer and Scherberger 2016). Decoding movement intentions from upstream brain areas such as PPC and PMv, instead of decoding individual finger movements from M1, may allow for more rapid and intuitive control of a grasp BMI (Andersen, Aflalo, and Kellis 2019). S1 processes incoming sensory feedback signals from the peripheral nervous system. While it is not thought to participate in grasp planning per se, it processes proprioceptive signals during movement (Goodman et al. 2019) and imagined somatosensations (Jafari et al. 2020), (Bashford et al. 2021), which could be used for a grasp BMI.

The grasp circuit was first identified in a non-human primate model (NHP), a network of cerebral pathways involved from visual object presentation to grasp execution. During an object manipulation task, neurons in anterior intraparietal cortex (AIP), a sub-region of PPC, and F5 (a PMv

analog in NHP) have both shown selectivity for object features, such as size, shape, and orientation (Akira Murata et al. 1997), (A. Murata et al. 2000), (Sakata 1995), (Taira 1998). In contrast to M1, information about the attempted grasp can already be decoded during cue presentation with high accuracy, implicating a role in motor planning (Carpaneto et al. 2011), (Townsend, Subasi, and Scherberger 2011), (Michaels and Scherberger 2017), (Schaffelhofer and Scherberger 2016). In human electrophysiological studies, some of these results have been replicated in AIP [-38 lateral, -53 posterior, 46 superior], demonstrating the ability to decode grasp planning and intention while a human participant performed motor imagery of one of five cued grasps (Klaes et al. 2015). However, it remains to be seen if an analogous relationship exists between pre-clinical results in NHP F5 and human PMv.

The supramarginal gyrus (SMG) has been hypothesized as a region specialized for complex tool use, evolutionarily evolving from a duplicate region of NHP AIP (Orban and Caruana 2014). Functional magnetic resonance imagining (fMRI) studies have shown activation of SMG during observed tool use, a finding which could not be replicated in presumed analogous anatomical regions of cortex in NHP (R. Peeters et al. 2009). Other studies confirmed SMG activity modulates during grasping and manipulation of objects (Sakata 1995), reaching (Filimon et al. 2009), and tool use (Gallivan et al. 2013), (Orban and Caruana 2014), (McDowell et al. 2018), (Buchwald, Przybylski, and Króliczak 2018), (Reynaud et al. 2019). Additionally, one study demonstrated SMG's involvement in both the planning and execution of (pantomimed) tool use (Johnson-Frey 2004b). Furthermore, different functional connectivity in left SMG has been demonstrated for tool use and tool transport gesturing (Garcea and Buxbaum 2019). These characteristics highlight SMG's rich potential as a source of grasp related neural signals in the human cortex, which could be used for BMI control of a prosthetic hand.

Recent studies in S1 indicate its potential as a target site for BMI applications in patients with tetraplegia. Human and NHP studies have demonstrated decoding of hand kinematics during executed hand gestures (Branco et al. 2017) and before contact during object grasping (Okorokova et al. 2020), respectively. These results suggest that neural signals could also be present during imagined movement (Zhang et al. 2017). Furthermore, modulation of S1 neurons during motor imagery of reaching (Jafari et al. 2020), as well as decoding of imagined cutaneous and

proprioceptive sensations (Bashford et al. 2021), has been demonstrated for the same participant whose data underlies this work. If grasping motor imagery can be robustly decoded, a single implant in S1 could allow for a bidirectional BMI, able to decode grasp intentions and utilize electrical stimulation to evoke somatosensations (Armenta Salas et al. 2018).

In this work, a tetraplegic participant performed motor imagery of several different grasps, while neurophysiological responses were captured from three implant sites using recording microelectrode arrays, the supramarginal gyrus (SMG), the ventral premotor cortex (PMv) and the primary sensory cortex (S1). We evaluated the decodability of these imagined grasps in the context of evaluating suitability for BMI applications. We hypothesized that grasp motor imagery would modulate activity in all three brain areas, while only the higher-level brain areas SMG and PMv would modulate during visual cue presentation.

Methods

Data and code availability

All analyzes were conducted in MATLAB using previously published methods and packages. MATLAB analyses scripts and preprocessed data are available on GitHub (https://doi.org/10.5281/zenodo.6330179).

Experimental model and subject details

A tetraplegic participant was recruited for an IRB- and FDA-approved clinical trial of a brain-machine interface, and he gave informed consent to participate. The participant suffered a spinal cord injury at cervical level C5 two years prior to participating in the study.

Implants

The targeted areas for implant were the left ventral premotor cortex (PMv), supramarginal gyrus (SMG), and primary somatosensory cortex (S1). Exact implant site within PPC and PMv was identified using fMRI while the participant performed imagined reaching and grasping tasks. The subject performed precision grip, power grip, or reaches without hand shaping of objects in different orientations (Aflalo, Kellis et al., 2015). For localization of the S1 implant, the subject was touched on areas with residual sensation on the biceps, forearm, and thenar eminence during fMRI,

and reported the number of touches (Armenta Salas et al. 2018). In November 2016, the participant underwent surgery to implant one 96-channel multi-electrode array (Neuroport Array, Blackrock Microsystems, Salt Lake City, UT) in SMG and PMv each, and two 7 x 7 sputtered iridium oxide film-tipped microelectrode arrays with 48 channels each in S1.

Data collection

Recording began two weeks after surgery and continued one to three times per week. Data for this work were collected between 2017 and 2019. Broadband electrical activity was recorded from the NeuroPort arrays using Neural Signal Processors (Blackrock Microsystems, Salt Lake City, UT). Analog signals were amplified, bandpass filtered (0.3-7500 Hz), and digitized at 30,000 samples/sec. To identify putative action potentials, these broadband data were bandpass filtered (250-5000 Hz), and thresholded at -4.5 the estimated root-mean-square voltage of the noise. Waveforms captured at these threshold crossings were then spike sorted by manually assigning each observation to a putative single neuron, and the rate of occurrence of each "unit", in spikes/sec, are the data underlying this work. Units with firing rate <1.5 Hz were excluded from all analyses. To allow for meaningful analysis of individual datasets, recording sessions where high levels of noise prevented us from isolating more than three units on an array were excluded. This resulted in the removal of three PMv datasets. The rounded average number of recorded units per session was 55 +/- 17 for SMG, 12 +/- 9 for PMv, and 119 +/- 48 for S1.

Experimental Task

We implemented a task that cued five different grasps with visual images taken from the *Human Grasping Database* (Feix et al. 2016) to examine the neural activity related to imagined grasps in SMG , PMv and S1. The grasps were selected to cover a range of different hand configurations and were labeled “Lateral”, “WritingTripod”, “MediumWrap”, “PalmarPinch”, and “Sphere3Finger” **(Figure 3-1A).**

Go task

Each trial consisted of four phases, referred to in this paper as ITI, cue, delay, and action **(Figure 3-1B).** The trial began with a brief inter-trial interval (2 sec), followed by a visual cue of one of the five specific grasps (4 sec). Then, after a delay period (gray circle onscreen; 2 sec), the participant was

instructed to imagine performing the cued grasp with his right (contralateral) hand (Go trials; green circle on screen; 4 sec). Three datasets had a longer action phase. For these, only data from the first four seconds of the action phase were included in the analysis.

Go/No-Go task

In a Go/No-Go version of this task, the participant was presented with either a green circle (Go condition) or a red circle (No-Go condition) after the delay, with instructions to imagine performing the cued grasp as normal during the Go condition (Go trials), and to do nothing for the No-Go condition (No-Go trials). In both variations of the task, conditions and grasp types were pseudo randomly interleaved and balanced with eight trials collected per combination **(Figure 3-1B)**.

Table 2-1 illustrates the number of recording sessions for each task variation.

The participant was situated 1 m in front of a LED screen (1190 mm screen diagonal), where the task was visualized. The task was implemented using the Psychophysics Toolbox (Brainard, 1997; Pelli,

Task Area	Go task	Go/No-Go task
SMG	6	9
PMV	6	6
S1	6	7

Table 0-1 | Number of recording sessions per task for each brain area.

1997; Kleiner et al, 2007) extension for MATLAB (MATLAB. (2018). 9.7.0.1190202 (R2019b). Natick, Massachusetts: The MathWorks Inc.).

Neural firing rates

Firing rates of sorted units were computed as the number of spikes that occurred in 50ms bins, divided by the bin width, and smoothed using a Gaussian filter with kernel width of 50ms to form an estimate of the instantaneous firing rates (spikes/sec). For the Go condition, 40 trials (8 repetitions of 5 grasps) were recorded per block. For the No-Go condition, two consecutive blocks of 40 trials (4 repetitions of 5 Go and 5 No-Go grasps) were recorded and combined, to accommodate the participant with shorter tasks.

Quantification and statistical analysis

All analyses were performed using MATLAB R2020b.

Linear regression analysis

To identify units that exhibited selective firing rate patterns (or tuning) for the different grasps, linear regression analysis was performed in two different ways: 1) step by step in 50ms time bins to allow assessing changes in neuronal tuning over the entire trial duration; 2) averaging the firing rate of specified time windows during the cue (1.5s) and action phase (2s), allowing to compare tuning between both phases. The model returns a fit that estimates the firing rate of a unit based on the following variables:

$$FR = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5$$

where FR corresponds to the firing rate of that unit, and β corresponds to the estimated regression coefficients. A 48 x 5 indicator variable, X, indicated which data corresponded to which grasp. The first 8 rows were the average firing rate of the ITI phase, and indicated the offset term $\beta_{o,}$ or baseline condition. These rows had only zeros. The next 40 rows indicated the trial data, for example, if the first trial was “Lateral” (grasp 1), it would have a 1 in column 1, and zeros in all other columns.

In this model, β symbolizes the change of firing rate from baseline for each grasp. A student’s t-test was performed to test the hypothesis of $\beta = 0$. A unit was defined as tuned if the hypothesis could be rejected ($p < 0.05$, t-statistic). This definition allows for tuning of a unit to zero, one, or multiple grasps during different time points of the trial.

Linear regression significance testing

To assess significance of unit tuning, a null dataset was created by repeating linear regression analysis 1000 times with shuffled labels. Then, different percentile levels of this obtained null distribution were computed and compared to the actual data. Data higher than the 95th percentile of the null - distribution was denoted with a * symbol, higher than 99th percentile was denoted with **, and higher than 99.9th percentile was denoted with ***.

Classification

Using the neuronal firing rates recorded in this task, a classifier was used to evaluate how well the set of grasps could be differentiated during each phase. For each session and each array individually, linear discriminant analysis (LDA) was performed, assuming an identical diagonal covariance matrix for data of each grasp. These assumptions, compared to a full diagonal covariance matrix, resulted in best classification accuracies. Classifiers were trained using averaged data from each phase, which were either 2s (ITI, delay) or 4s (cue, action). We applied principal component analysis (PCA) and selected the 10 highest principal components (PCs), or PCs explaining more than 90% of the variance (whichever was higher), for feature selection on the training set. When less than 10 PCs were available, all features were used. This feature selection method allowed us to compare if there was a correlation between the number of tuned units and classification accuracy, without selecting tuned units as features. The unit yield in PMv was generally lower than in SMG and S1; however, significant classification accuracies were still obtained with a limited number of features. Between 12 and 21 PCs were used in SMG, 6 and 16 in PMv, and 18 and 27 in S1. Leave-one-out cross-validation was performed to estimate decoding performance. A 95% confidence interval was computed by the student's t-inverse cumulative distribution function.

Classification performance significance testing

To assess the significance of classification performance, a null dataset was created by repeating classification 1000 times with shuffled labels. Then, different percentile levels of this null distribution were computed and compared to the mean of the actual data. Mean classification performances higher than the 95th percentile were denoted with a * symbol, higher than 99th percentile were denoted with **, and higher than 99.9th percentile were denoted with ***.

Neuron dropping curve and cross-phase classification

The neuron dropping curve represents the evolution of the classification accuracy based on the number of neurons used to train and test the model. All available neurons were used for all brain areas. Cross-phase classification was performed to investigate how well a model trained on data of the cue phase can predict data of the action phase, and vice-versa. Classification with eightfold cross validation was performed for each subset of neurons selected for classification. First, one of the neurons was randomly selected, and the classification accuracy on the cue and action phase was

computed with a model trained on either the action phase or the cue phase. Then, a new subset of two random neurons was selected, and classification accuracy was again computed. This was performed until all available neurons were randomly added. PCA was performed on the dataset. To avoid overfitting by using more features than observations (40), the maximum number of principal components used was 20, and the process was repeated 100 times. The prediction accuracy was averaged over the cross-validation folds, and the mean with 95% confidence interval (bootstrapped) was plotted against the number of neurons.

Results

Grasp representation in SMG, PMv, and S1 was characterized by decoding five imagined grasps, cued with visual images taken from the *Human Grasping Database* (Feix et al. 2016). We evaluated the brain regions' potential for a grasp BMI in two ways; firstly, by quantifying grasp tuning in the neuronal population and secondly, by assessing how well individual grasps were decodable from each area. SMG, PMv, and S1 neural populations showed significant grasp selectivity, making them candidates for grasp BMI implantation sites. Additionally, if large enough neuronal populations are present, both SMG and PMv show high grasp selectivity, making them noteworthy candidates for grasp BMI implantation sites.

Motor imagery task design

The motor imagery task contained four phases: an inter-trial interval (ITI), a cue phase, a delay phase, and an action phase (**Figure 3-1A**). The Go variation of the task consisted of only Go-trials, with performed motor imagery during the action phase. A Go/No-Go variation of the task contained an action phase with randomly intermixed Go trials and No-Go trials. This control condition verified the participant could control motor imagery-related activity at will.

Go trial results were quantitatively similar in both the Go and Go/No-Go variations of the task, as assessed through a t-test between classification accuracies ($p > 0.05$ for all). Therefore, neurons involved during Go-trials in both tasks were pooled over all session days (see **Table 1**), resulting in 819 SMG Go task units, 504 SMG No-Go task units, 146 PMv Go task units, 78 PMv No-Go task units, 1551 S1 task Go units, and 948 S1 No-Go task units.

SMG, PMv, and S1 show significant tuning to grasps during motor imagery

Smoothed firing rates of example units for SMG and PMv during the Go/No-Go version of the task displayed neuronal modulation to the grasp "Sphere3Finger" in **Figure 3-1B**. Motor imagery evoked a strong response during the action phase of Go trials compared to the action phase of No-Go trials, where firing rate decreased back to baseline activity.

After establishing individual neural firing rate modulation during motor imagery for different grasps, we quantified the entire neuronal population's selectivity for each grasp. To compare selective neural activity within task epochs (image cue, Go-task action phase, No-Go task action phase), we determined the duration of selective (or tuned) activity of the neural population during each phase. Tuning of a neuron to a grasp was determined by fitting a linear regression model to the firing rate in 50ms time bins (see methods).

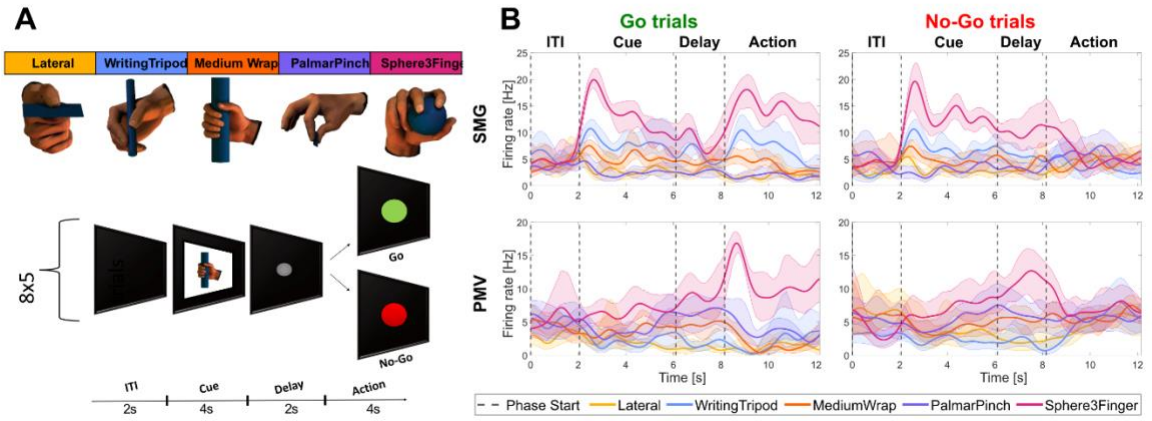

Figure 0-1 | Neurons in the cortical grasp circuit encode grasp types. A) Grasp images from the *Human Grasping Database* (Feix et al., 2016) were used to cue motor imagery in a tetraplegic human. The task was composed of an inter-trial interval (ITI), a cue phase displaying one of the grasp images, a delay phase, and an action phase. The action phase was composed of intermixed Go trials (green), during which the participant performed motor imagery and No-Go trials (red), during which the participant rested. **B)** Example smoothed firing rates of neurons in SMG and PMv during Go (left) and No-Go (right) trials. The plots show the smoothed average firing rate of two example units (solid line, shaded area 95% bootstrapped confidence interval) for 8 trials of each grasp, with vertical lines representing the beginning of each phase.

Population analysis (**Figure 3-2A**) of Go trials revealed two main peaks of activation in SMG and PMv, one at cue presentation (54.8% SMG, 41.1% PMv) and another during the action phase (37.4% SMG, 39.0% PMv). For S1, only a minor increase in neural tuning was observed during the action phase. During No-Go trials, neuronal activity decreased around 1s after start of the action phase (**Figure 3-2B,** action phase). This pattern could indicate the formation of a motor plan during the cue phase, and a brief period of activity during the action phase when this plan was canceled.

The peaks of activity were selected to compute individual grasp tuning. Time windows incorporating the peaks began 250ms after the start of either the cue or action phase (to account for processing latencies), and were respectively 1.5s and 2s long (gray lines, top of **Figure 3-1C,D**). A longer time window was chosen for the action phase, as the exact onset of motor imagery is not possible to measure.

To assess if grasp tuning was significant, results were compared to a shuffled condition, where grasp labels were randomly reassigned (see methods). As linear regression uses the ITI phase as a baseline condition, shuffled results were proportional to the general increase of activity in the neuronal population. Tuning was significant during the Go-trial peak activity for all brain areas (**Figure 3-2C**). As expected, tuning was not significant in the ITI condition. During the cue phase, results were significant in SMG and PMv, but not significant in S1. During the action phase, no significance was found during No-Go trials for all brain areas (**Figure 3-2D**). These results highlight grasp-dependent neuronal activity during cue presentation in SMG and PMv, and during instructed motor imagery in all brain areas.

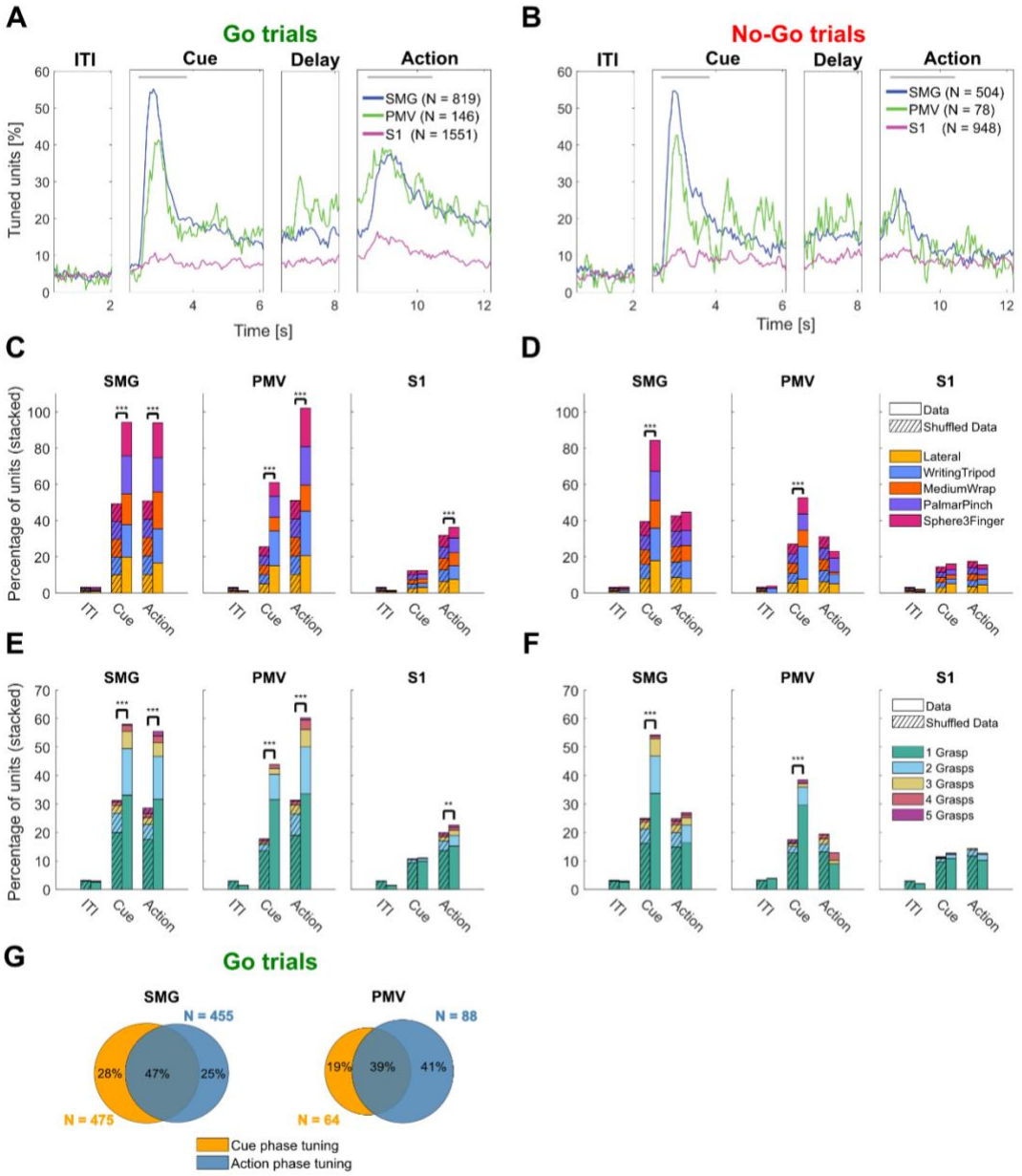

Figure 0-2 | Population analysis of grasp tuning. A) Percentage of tuned units to grasps for Go trials in 50ms time bins in SMG, PMv and S1, over the trial duration. The gray lines represent cue and action analysis windows for Figures B,C and D. **B)** Same as A) for NoGo trials. **C)** Stacked percentage of units tuned for each grasp in ITI, cue phase and action phase window during Go trials. Significance was calculated by comparing data (right bar) to a shuffle distribution (striped lines, left bar). **D)** Same as C) for No-Go trials. **E)** Stacked percentage of units tuned to one, two, three, four, and five grasps during the cue phase and the action phase analysis window during Go trials. Significance was calculated as described previously. **F)** Same as E) for No-Go trials. **G)** Overlap of tuned units between cue and action analysis window during Go trials for SMG and PMv.

How many different grasps were individual units able to represent? Results were consistent across all brain areas with most tuned units selective to one grasp (**Figure 3-2E**). In SMG and PMv, additional units were tuned to multiple grasps, demonstrating mixed grasp encoding within the population. As before, results were significant during cue presentation in SMG and PMv, during Go-trial action phase in all brain areas, but not during the ITI, or during the No-Go trial action phase (**Figure 3-2E,F**).

Similar to previous analysis methodologies (Akira Murata et al. 1997), (Sakata 1995), (Taira 1998), (Klaes et al. 2015), we separated tuned units into three categories: those tuned during the cue phase ("visual units"), those tuned during Go-trial action phase ("motor-imagery units"), and those tuned during both ("visuo-motor units"). All three neuron types were found in SMG and PMv **(Figure 3-1B, Figure 3-2G)**.

SMG, PMv, and S1 show significant classification accuracy during grasp motor imagery

To assess each brain region's potential use for BMI applications, we evaluated decodability of individual imagined grasps using linear discriminant analysis (LDA; see Methods). Significant motor imagery decoding was observed in all brain areas (**Figure 3-3A)**. Black dots indicate individual session results; red dots indicate averaged shuffled results. The mean +/- 95% confidence interval (c.i.) was computed over individual sessions. Significant classification accuracies were obtained for cue, delay and Go-action phases in SMG ($p < 0.001$), cue ($p < 0.01$), delay ($p < 0.05$), Go-action phases ($p < 0.001$) in PMv, and Go-action phase ($p < 0.5$) in S1. For No-Go trials, significant classification accuracies were obtained in the cue and delay phase for SMG ($p < 0.001$, $p < 0.01$), and the cue phase in PMv ($p < 0.05$), but not the action phase (**Figure 3-3B**). Importantly, these results mirror the findings in **Figure 3-2C,E**, indicating that significant grasp tuning can predict significant classification accuracies. A confusion matrix averaged over all sessions of Go-trials in SMG and PMv during the action phase suggests that all grasps can be decoded (**Figure 3-3C,D)**.

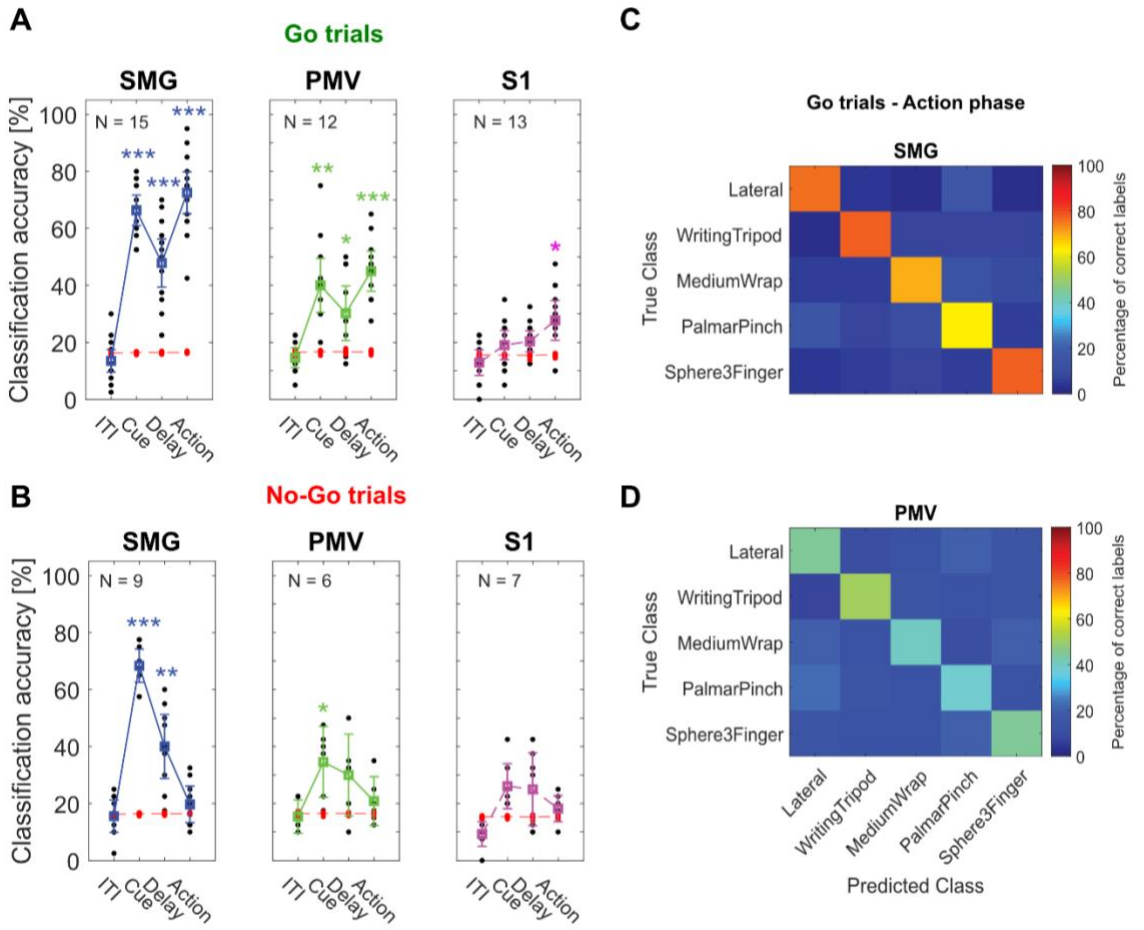

Figure 0-3 | Significantly decodable grasps from all brain areas during motor imagery. A) Classification was performed for each session day individually using leave-one-out cross-validation (black dots). PCA was performed. 95% c.i.s for the session means were computed. Significance was evaluated by comparing actual data results to a shuffle distribution (averaged shuffle results = red dots, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$). **B)** Same as A) for No-Go trials. **C)** Error matrix during Go-trial action phase for SMG, averaged over all session days. **D)** Same as C) for PMv.

SMG and PMv show high generalizability of grasp encoding in the neural population

We addressed generalizability of grasp encoding in the neural population via two analyses: cross-phase classification and stability across different population sizes.

Cross-phase classification examined the similarities of neural processes across the cue and action phases (see methods). We trained a classification model on a subset of the data of one phase (e.g. cue phase), and tested it on two different subsets taken from the cue and action phases. If a model trained on the cue phase does not generalize to the action phase, distinct neural processes might

be present in each phase. However, if the model generalizes well, common cognitive processes might be occurring in both phases. In parallel, a neuron dropping analysis tracked the evolving classification accuracy as units were removed or added to the pool of predictors (see methods). The analysis was performed separately for each of the implanted brain regions, with 100 repetitions of eight-fold cross-validation.

Results were averaged over 8-folds and bootstrapped confidence intervals (c.i.s.) of the mean were computed over 100 repetitions (**Figure 3-4)**. Stable results led to small c.i.s, ranging from ±2.88% to ±0.05% for SMG, ±2.83% to ±0.54% for PMv, and ±2.36% to ±0.8% for S1, decreasing with the number of available units. SMG and PMv showed strong shared activity between the cue and action phases. When training on the cue phase, and testing on the cue and action phases, we observed good generalization of the model in SMG, with overlapping c.i.s, diverging only at high unit counts. In PMv, the generalization was lower, but showed similar trends, while decoding remained at chance level for S1 (**Figure 3-4 A,B,C** Train: Cue Phase). However, when training on the action phase, and evaluating on the cue phase, lower generalization of the model was observed in SMG and PMv **(Figure 3-4 A,B,C** Train: Action Phase).

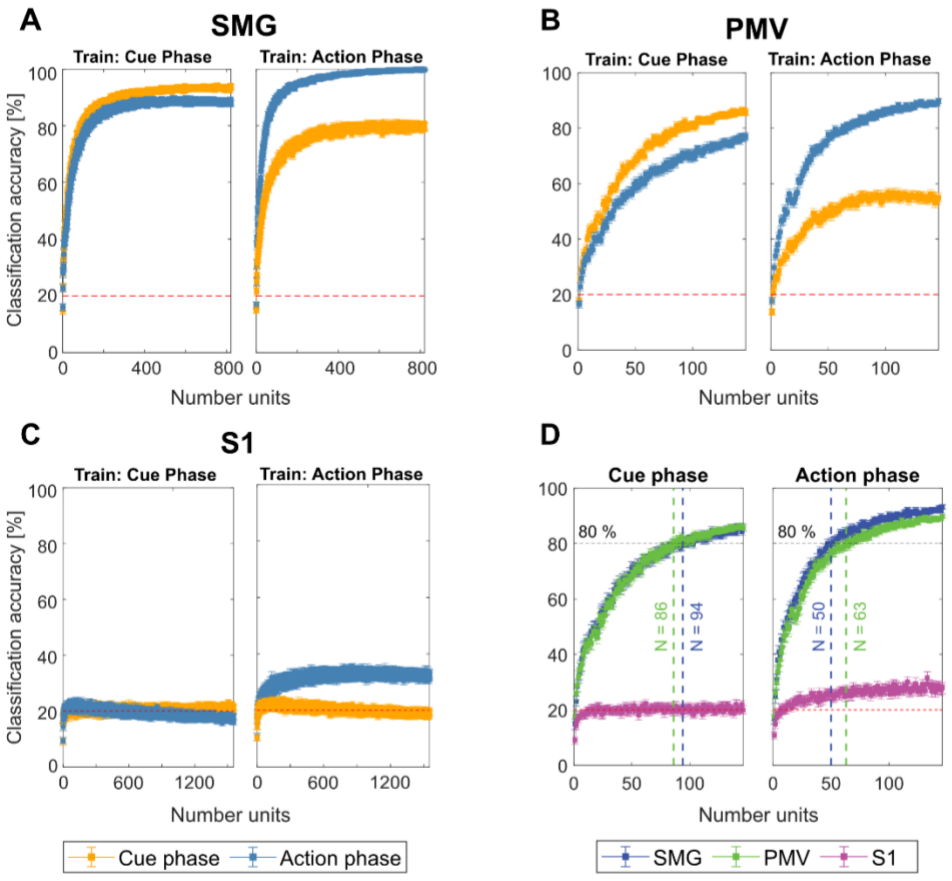

Figure 0-4 | SMG and PMv show high generalizability of grasp encoding in neuronal populations. A-C) Neuron dropping curves were performed in SMG, PMv and S1 over 100 repetitions of eight-fold cross validation. The first 20 PC's were used as features. The model was trained once on the cue phase and applied on both cue and action phases (Train: Cue phase), and vice-versa (Train: Action phase). The mean classification accuracy with bootstrapped 95% c.i.s are plotted. **D)** The first 140 units of each brain area were plotted together to compare the number of units required for 80% classification accuracy. SMG and PMv results were similar, with less units needed for classification during the action phase compared to the cue phase.

During the action phase, SMG peaked at 99% decoding accuracy when all recorded units were included in the analysis (**Figure 3-4A**). In S1, decoding accuracy during the action phase peaked around 32%, even when the pool of available neurons increased (**Figure 3-4C**). As PMv did not reach its peak decoding accuracy due to fewer number of units recorded (**Figure 3-4B**), performance of SMG and PMv at the same population levels was compared directly. **Figure 3-4D** depicts the number of features needed to obtain 80% classification accuracy during cue (left) and action (right) phase. During the cue phase, 94 units in SMG and 86 units in PMv were needed.

During the action phase, 80% classification accuracy was obtained with 50 units in SMG, and 63 units in PMv. These results demonstrate SMG's and PMv's potential for comparable grasp decoding. If higher neuronal population counts were available, excellent grasp classification results can be expected in both brain areas.

PMv had a limited number of neurons available for each daily session. It is possible that some units were included multiple times across multiple days, potentially reducing the amount of independent available information. However, since the highest classification accuracy during the action phase was higher for the neuron dropping curve (90%) than for individual session days (65%, **Figure 3-2A**), new grasp information was available by combining units across several days.

Discussion

In this work, we demonstrated that motor imagery of five unique grasps was well represented by the firing rates of neuronal populations, and could be decoded significantly above chance level in the supramarginal gyrus (SMG), the ventral premotor cortex (PMv), and the primary sensory cortex (S1). SMG and PMv encoded grasp information both during cue presentation and during motor imagery with similar neuronal activity patterns. Equal numbers of units in the neuronal populations of SMG and PMv showed comparably excellent grasp encoding capabilities, demonstrating high potential for grasp BMI applications in both areas.

To demonstrate the participant had volitional control of motor imagery during the action phase, and observed activity was not due to some external factor, interleaved No-Go trials served as a control. During No-Go trials in the action phase, unit tuning was not significantly different from a shuffled distribution (**Figure 3-2D,F**), and classification was not significantly different from chance (**Figure 3-3A**). A non-significant peak in tuning was observed in **Figure 3-2B** (No-Go action phase trials), potentially indicating the formation of a motor plan before the No-Go cue that then dissipated in the action phase. Similar cancelled plans have been previously observed in PPC of NHPs for reach and saccade plans (Cui and Andersen 2007).

S1 encodes imagined grasps significantly, but does not improve with population size

While S1 grasp motor imagery classification was significant (**Figure 3-3A**), performance did not improve with increased population sizes as was seen with SMG and PMv (**Figure 3-4D**). This could be an indication of limited grasp information within the S1 population, or highly correlated firing units. Firstly, no actual movement was performed, likely decreasing the occurrence of proprioceptive signals (limited available information). Secondly, the task design might have only weakly engaged the neural populations we recorded from, as the electrode implant mostly covered the contralateral arm area (Armenta Salas et al. 2018). A different task, that involved the arm by reaching to grasp an object, may have elicited stronger neuronal activity (Jafari et al. 2020). Thirdly, units in S1 showed mostly grasp independent increases in activity compared to baseline (**Figure 3-2C,E**), possibly indicating that the grasp-related responses were not different enough to support stronger decoding in S1 (correlated information).

SMG and PMv show significant grasp activity during the visual cue and motor imagery

SMG's cue phase activity rose faster, and peaked higher compared to activity during motor imagery (**Figure 3-2A**). A study showed grasp planning in SMG was disrupted by TMS as early as 17ms after cue presentation, suggesting a causal role in grasp planning and execution (Potok et al. 2019).

Evidence for mixed visual and motor activity during the action phase

While human participants can self-report strategies employed while performing internal cognitive tasks, cue processing and motor imagery do not have independently observable behavioral outputs to correlate with the measured neural data. Our analysis showed generalizable representation **(Figure 3-4A,B**) and overlapping tuning (**Figure 3-2G**) in both SMG and PMv during both the cue and action phases. Multiple explanations for generalized neural activity observed during these tasks are plausible. During cue presentation, an increase in neural activity could represent visual feature extraction of the presented cue (visual processes). Alternatively, activity could be independent of visual input and represent planning activity of the cued grasp (motor processes). Additionally, activity could be related to memory or semantic meaning, as the participant remembers the instructed grasp (cognitive processes). Finally, a combination of all these processes might be at play. While proving a definitive answer to these questions is beyond the scope of this paper, performing

cross-phase classification between the cue and action can help identify similar or distinct cognitive processes within the observed data.

Cross-phase classification found similar neuronal activity in the cue and action phases in both SMG and PMv **(Figure 3-4A,B**). This agrees with our finding of overlapping neuronal populations tuned during both the cue and action phase (**Figure 3-2G**). One explanation for these similarities could be that the participant is performing "visual imagery" rather than motor imagery during the action phase, by recalling a mental image of the grasp (**Figure 3-4A,B** Train: Cue Phase). Cue phase activity can partly be explained during the action phase (classification performance 80% SMG, 55% PMv) (**Figure 3-4A,B** Train: Action Phase), but neuronal activity unique to the action phase exists (classification performance 99% SMG, 89% PMv). This generalization from the cue to action phase is not bidirectional (from action to cue phase). Furthermore, training a classifier on neural data from spoken grasps during the action phase did not generalize to neural data from motor imagery during the action phase **(Figure 3-5C**), while partly generalizing with neural data during cue phase **(Figure 3-S1A)**. Therefore, we argue this additional information during the action phase is likely motor-related and thus fundamentally differs from neural activity during the cue phase.

Good generalization of the model to both cue and action phases when training on the cue phase could indicate motor components as well as visual components. PMv has been shown to represent planning activity of the grasp in NHP experiments (Schaffelhofer and Scherberger 2016). Therefore, planned hand shape as well as visual object features can modulate neuronal firing rates within the cortical grasp circuit during a grasp task. In SMG, a fMRI study demonstrated planning activity for grasping tools that were previously manipulated without vision, hinting that SMG's cue phase activity is likely not to be only visual (Styrkowiec, Nowik, and Króliczak 2019).

Cue phase activity could represent semantic or memory processing, i.e., the abstract concept of each cued grasp. During tool use, SMG is hypothesized to integrate the appropriate grasp type with the knowledge of how to use the tool (Osiurak and Badets 2016; Vingerhoets 2014), which requires access to semantic information. As our current task design does not allow the differentiations of these cognitive processes, further experimentation is necessary. For instance, cueing grasps with non-visual sensory cues and observing if cue phase activity is still present, might allow the

dissociation between visual, motor, and semantic processes, and help clarify the roles of SMG and PMv in the human grasp circuit.

When analyzing SMG and PMv for potential grasp BMI applications, both performed similarly. While SMG displays stronger encoding of grasps than PMv on a session-to-session basis (**Figure 3-3A**), these results are likely due to the small number of units we were able to record from the PMv array on individual days. The neuron dropping analysis illustrates that when identical neuronal population are present, SMG and PMv have similar grasp decoding abilities (**Figure 3-4D**).

Conclusion

In this study, we demonstrate grasps are well represented by single unit firing rates of neuronal populations in human SMG and PMv during cue presentation. During motor imagery, individual grasps could be significantly decoded in all brain areas. SMG and PMv achieved similar highly-significant decoding performances, demonstrating their viability for a grasp BMI.

Acknowledgements

We wish to thank L. Bashford, H. Jo, and I. Rosenthal for helpful discussions and data collection. We wish to thank our study participant FG for his dedication to the study which made this work possible. This research was supported by the NIH National Institute of Neurological Disorders and Stroke Grant U01: U01NS098975 (S.K.W., S.K., D.B., K.P., C.L. and R.A.A.) and by the T&C Chen Brain-machine Interface center (S.K.W., D.B., R.A.A.).

Author contributions

S.K., S.K.W., and R.A.A. designed the study. S.K.W. and S.K. developed the experimental tasks. S.K.W., S.K., and D.B. analyzed the results. S.K.W., S.K., D.B. and R.A.A. interpreted the results and wrote the paper. K.P. coordinated regulatory requirements of clinical trials. C.L. and B.L. performed the surgery to implant the recording arrays.

Chapter 4:

Grasp concept encoding through different sensory modalities in human posterior parietal cortex

The following chapter's contents are taken and adapted from a conference talk given at the 8th international BCI meeting, 2021, with modifications done to fit the dissertation format.

S.K. Wandelt, S. Kellis, L. Bashford, B. Lee, C. Li, R. A. Andersen (2021) Grasp concept encoding through different sensory modalities in human posterior parietal cortex, talk, 8th international BCI meeting, 2021.

Abstract

Grasping and manipulation of objects are important aspects of human independence and represent critical losses in paralysis due to spinal cord injury (SCI). Intracortical recordings from posterior parietal cortex (PPC) and the ventral premotor cortex (PMv) in participants with tetraplegia have previously been shown to exhibit planning and execution activity during motor imagery of different grasp shapes using visual images of grasps or objects. However, areas of the cortical grasp circuit are also involved in a variety of other cognitive tasks, such as visual word recognition and phonological processing. These different sensory and behavioral paradigms could potentially modulate preparatory activity and how grasps are represented during motor imagery. To understand how cue modalities affect motor imagery in different brain areas of the cortical grasp circuit, we instructed two tetraplegic participants to perform a grasp motor imagery task. Participant 1 had implants in the supramarginal gyrus (SMG), and PMv and participant 2 had an implant in the anterior intraparietal cortex (AIP). We tested image, auditory, and written cue modalities and evaluated how grasp-related information was represented in different phases of the task. We found cue-independent grasp motor imagery, suggesting the underlying meaning of the grasp concept remained the same. During the cue phase, modulation to different sensory modalities depended on the observed brain area, with SMG showing cue-modality dependent activity for image, audio, and written cues. Generalization between auditory and written cue modality was higher than for the image modality, suggesting SMG represents language processes. To summarize, we show that certain areas of PPC can performs visual, auditory, and written cue integration, as well as motor planning activity, while performing a grasp motor imagery task.

Introduction

Brain-machine interfaces (BMIs) have emerged as a promising tool for restoring lost motor functions in individuals with paralysis. Of particular interest for grasp motor imagery applications are intracortical BMI implants located in the human cortical grasp circuit. Neuronal signals recorded from these brain areas have allowed for high-accuracy decoding of reach and grasp movements (Collinger et al. 2013; Klaes et al. 2015; Tyson Aflalo et al. 2015; Andersen, Aflalo, and Kellis 2019). In recent years, the posterior parietal cortex (PPC) and the ventral premotor cortex (PMv) within the cortical grasp circuit have been explored for grasp motor imagery (Andersen, Aflalo, and Kellis 2019; Wandelt et al. 2022b; Schaffelhofer, Agudelo-Toro, and Scherberger 2015b; Meng Wang et al. 2020; Klaes et al. 2015). These higher-level brain areas are thought to play critical roles in sensorimotor transformations that allow for flexible and adaptive grasp control, and could allow for improved BMI applications (Andersen and Buneo 2002; Andersen et al. 2014a; Cui and Andersen 2007; Andersen, Aflalo, and Kellis 2019; Gallego, Makin, and McDougle 2022).

Standard motor imagery experiments are composed of a cueing phase during which the participant gets instructed which grasp to perform, followed by a delay phase, and an action phase, during which motor imagery is being performed. The employed cued is often visual, depicting a hand position, an object, or an abstract image (Schaffelhofer and Scherberger 2016; Klaes et al. 2015; Wandelt et al. 2022b; Schaffelhofer, Agudelo-Toro, and Scherberger 2015b). Both PPC and PMv increase their firing rate already during cue presentation and motor planning, distinguishing them from downstream primary motor cortex (Wandelt et al. 2022b; Schaffelhofer and Scherberger 2016; Klaes et al. 2015).

Although these brain areas are known to play a critical role in grasp planning and execution in both human and non-human primates (NHP) (Wandelt et al. 2022a; Klaes et al. 2015; Schaffelhofer and Scherberger 2016; Townsend, Subasi, and Scherberger 2011), they also show involvement in a variety of other cognitive tasks, such as memory processing, written word recognition and phonetics (Rutishauser et al. 2018; T. Aflalo et al. 2020; Stoeckel et al. 2009; Sliwinska et al. 2012a; Oberhuber et al. 2016). Therefore, which cognitive process they represent during grasp preparation in unclear. Indeed, when cueing a participant with a visual image of a grasp or an object, several cognitive

processes could be at play. The activity could be visual in nature, as grasping requires extracting visual features of an object for appropriate grasp shaping. The activity could also be motor and joint position related, as the participant is planning a motor imagery grasp. Moreover, when the participant recognizes the grasp, memory and semantic processing are occurring. Finally, higher-level brain areas could also represent a mix of these processes at the same time. Therefore, different sensory and behavioral paradigms could potentially modulate preparatory activity and how grasps are represented during motor imagery in the cortical grasp circuit.

To understand the cognitive process represented during grasp cueing and planning on the neuronal level in human PPC and PMv, and to study impact of cue modality on motor imagery, we designed a grasp motor imagery task using three different sensory cues: a visual image cue, an auditory cue, and a written cue. These parameters allowed us to ask several questions. Firstly, which cognitive processes drives neural activity in PPC and PMv? Secondly, how similar are sensory modalities processed in different regions of the cortical grasp circuit? Lastly, is motor imagery representation invariable to cue modality? For instance, it is unclear if non-visual cues would allow a participant to transmit the same precise and detailed information required for fine grasp motor control as visual images of a grasp.

In our experiment, two tetraplegic patients performed grasp motor imagery of five different grasps using one of three sensory cue modalities. The first patient had intracortical implants in the supramarginal gyrus (SMG) and PMv, and the second participant had intracortical implants in the anterior intraparietal cortex (AIP).

We found that regardless of the employed cue modality, motor imagery was decoded equally well. During the cue phase, while AIP and PMv preferentially encoded the visual image cue, SMG showed unique neural modulation to all cues. Generalization between the auditory and written cue modality was higher than for the image modality, suggesting SMG may represent language processes. Understanding which cognitive processes are represented in areas of the cortical grasp circuit could allow tailoring cues to individual users' preferences and guide new applications for BMI devices.

Methods

For "Data and code availability", "Experimental model and subject details", "Method details", "Data collection" and "Quantification and statistical analysis" paragraphs, see Chapter 3, as methods were identical. This methods section only contains novel methods and experimental setups that have not been described in Chapter 3.

Subjects and implants

Two tetraplegic subjects were recruited for an IRB- and FDA-approved clinical trial of a brain-machine interface and gave informed consent to participate. The first subject, s2, suffered a spinal cord injury at cervical level C5 two years prior to participating in the study. The targeted areas for implant were the supramarginal gyrus (SMG), the ventral premotor cortex (PMv), and the primary somatosensory cortex (S1). In this work, only the SMG and PMv array were of interest. To identify exact implant sites within these regions, the subject performed imagined reaching and grasping tasks during functional magnetic resonance imaging (fMRI), see (Armenta Salas et al. 2018) for additional details. In 2015, s2 underwent surgery to implant one 96-channel multi-electrode array (Neuroport Array, Blackrock Microsystems, Salt Lake City, UT) in each of these areas.

Participant N1 suffered a spinal cord at level C4-C5 approximately two years before this study. He has residual movements in his upper arms but cannot move or feel his hands (Guan et al. 2022).

Data collection

Data for this work were collected between January and April 2020 for participant s2. For participant n1, data were collected between November 2021 and May 2022. Broadband electrical activity was recorded from the NeuroPort arrays using Neural Signal Processors (Blackrock Microsystems, Salt Lake City, UT). Analog signals were amplified, bandpass filtered (0.3-7500 Hz), and digitized at 30,000 samples/sec. To identify putative action potentials, these broadband data were bandpass filtered (250-5000 Hz), and thresholded at -4.5 the estimated root-mean-square voltage. Waveforms captured at these threshold crossing were spike sorted by manually assigning each observation to a putative single neuron. On each session day, all three tasks were recorded, and spikes were sorted in combination. This allowed observation of the behavior of each unit in the three different tasks.

Experimental Tasks

We implemented three tasks that evaluated the effect of different cue modalities on motor imagery in areas of the human cortical grasp circuit. The "Image" task cued five different grasps with visual images taken from the *Human Grasping Database* (Feix et al. 2016) to examine the neural activity related to imagined grasps in SMG , PMv, and AIP. The grasps were selected to cover a range of different hand configurations and were labeled "Lateral", "WritingTripod", "MediumWrap", "PalmarPinch", and "Sphere3Finger" **(Figure 4-1A)**. The "Auditory" tasks cued the participant with the auditory name of the grasp, while the "Written" task cued the participant with the written name of the grasp on the screen.

Image task:

Each trial consisted of four phases, referred to in this paper as ITI, cue, delay, and action**.** The trial began with a brief inter-trial interval (2 sec), followed by a visual cue of one of the five specific grasps (2 sec). Then, after a delay period (gray circle onscreen; 2 sec), the participant was instructed to imagine performing the cued grasp with his right (contralateral) hand (Go trials; green circle on screen; 4 sec).

Auditory task

An auditory variation of the task was constructed with the same task design outline above. During the cue phase, the visual image was replaced with the sound of the name of the grasp. Importantly, no visual information was shown on the screen during the cue phase.

Written task

In the written version of the task, the participants were cued with grasp name in written white text on black background.

On each of seven session days, a "Visual task", a "Auditory task" and a "Written task" were performed to allow comparisons between tasks.

The participants were situated 1 m in front of a LED screen (1190 mm screen diagonal), where the task was visualized. The tasks were implemented using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) extension for MATLAB (MATLAB. (2018). 9.7.0.1190202 (R2019b). Natick, Massachusetts: The MathWorks Inc.).

Cross-task classification

To evaluate the similarity of neuronal firing in the "Visual cue", the "Auditory cue" and the "Written cue" tasks, cross-task classification was performed. This method consisted of training a classifier on the averaged neuronal firing rates recorded during one of the tasks (e.g., "Visual cue") and evaluating it on the neuronal firing rates of all three tasks. A LDA with PCA and Leave-one-out cross validation was performed for each individual phase (see Methods section "Classification" in Chapter 3).

Results

To characterize which cognitive processes are occurring during grasp cueing and preparation in brain areas SMG, PMv, and AIP, and to understand their effect on grasp motor imagery, three different tasks were designed. The task varied the sensory modality employed for grasp cueing. The "Visual cue" task showed the visual image of the grasp, the "Auditory cue" task required listening to the audio sound of the grasp name, and the "Written cue" task required reading the written name of the grasp on the screen (**Figure 4-1**). The visual image did not require knowledge of the grasp name, in contrast to the auditory and the written cue. To ensure performance was not affected by not knowing which motor imagery to perform, and that the participants knew the correct names, a quiz was performed at the beginning of each session.

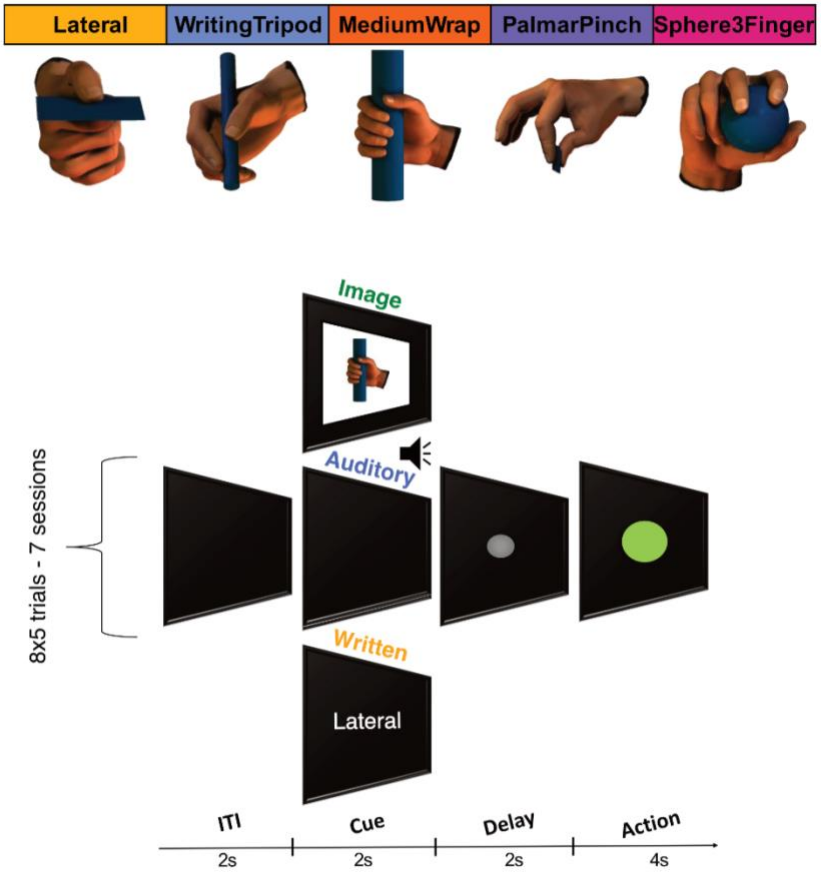

Figure 0-1 | Task design. Three tasks was designed to study the effect of different sensory cue modalities on grasp motor imagery. Tasks were composed of an inter-trial interval (ITI), a cue phase displaying one of three cue modalities, a delay phase, and an action phase. During the action phase, the participant was instructed to perform motor imagery of the instructed grasp. The Image cue tasked used visual images of a grasp as a cue (see Chapter 1). The Auditory cue task used the auditory sound of the name of the grasp. During the Auditory cue, the screen remained dark. For the Written cue task, the white name of the instructed grasp on black background appeared on the screen.

PMv and AIP preferentially encode visual cues

During each of seven session days, a "Visual cue", an "Auditory cue", and a "Written cue" task were run. Data were spike sorted, resulting in a total of 173 SMG units, 86 PMv units, and 606 AIP units for each task. First, tuning of units to different task parameters was computed using a linear regression analysis. The percentage of tuned units over trial duration for each of the tasks and brain regions was plotted in 50ms time bins (**Figure 4-2**). During the action phase, temporal dynamics for each cue modality were similar, suggesting sensorimotor transformation occurred regardless of the employed cue. Due to software constraints, the auditory cue was played ~250 ms later than the

image and written cue appeared. Therefore, to compare activity during the cue phase for the "Auditory task", neural data was aligned to the start of the cue phase. We observed differences in the strength of activity depending on the investigated brain area. While AIP and PMv neuronal populations appeared to modulate fasted and strongest to the image cue condition, neurons in SMG were the most engaged during the written cue. These results suggest areas in PPC and PMv may represent different cognitive processes during grasp planning.

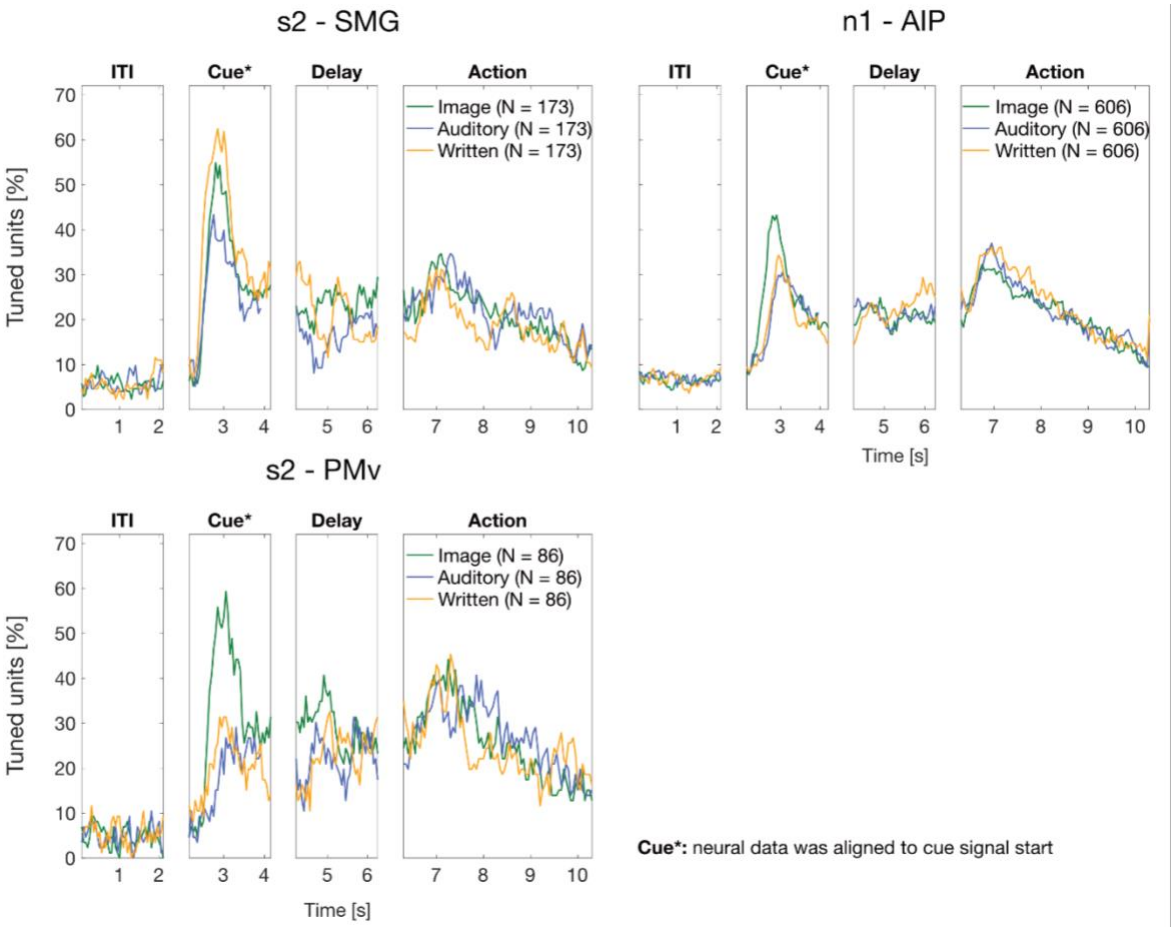

Figure 0-2 | Neuronal population dynamics to sensory modalities differ per brain area. The percentage of tuned channels was computed using linear regression analysis and plotted over the task duration in 50ms time bins for the "Image cue" task (green), "Auditory cue" task (blue), and "Written cue" task (yellow). In total, 173 SMG, 86 PMV, and 606 AIP units were obtained. While PMv and AIP show preferential modulation to the image cue, SMG had the highest percentage of tuned units to the written cue condition.

Next, we evaluated how well each brain region could decode five different grasps for each of the different grasp cues, and in each of the task phases. For each session day, leave-one-out cross-

validation was performed, PCA was performed on the training dataset, and PCs explaining 90% of the variance were kept. Average session results with 95% c.i. were plotted, and significance was evaluated by comparing results to a shuffled distribution. We found significant grasp motor imagery decoding in SMG, PMv, and AIP ($p < 0.01$), while results during the cue phase varied. Significant classification of grasps was possible for all cue modalities in SMG and AIP ($p < 0.01$). In PMv, decoding between grasp audio sounds was not significant, and slightly significant for written grasp words ($p < 0.05$). (**Figure 4-3A**).

We evaluated if classification accuracies in the cue and action phase were significantly different between data from different task modalities with a two-sample t-test (**Figure 4-3B**). Results mirrored the finding of the tuning analysis. PMv and AIP showed significantly higher decoding of grasp images than of grasp audio sounds and written grasp names. In SMG, decoding of written grasp words elicited significantly higher decoding accuracies. During motor imagery, classification accuracies were not significantly different for each task.

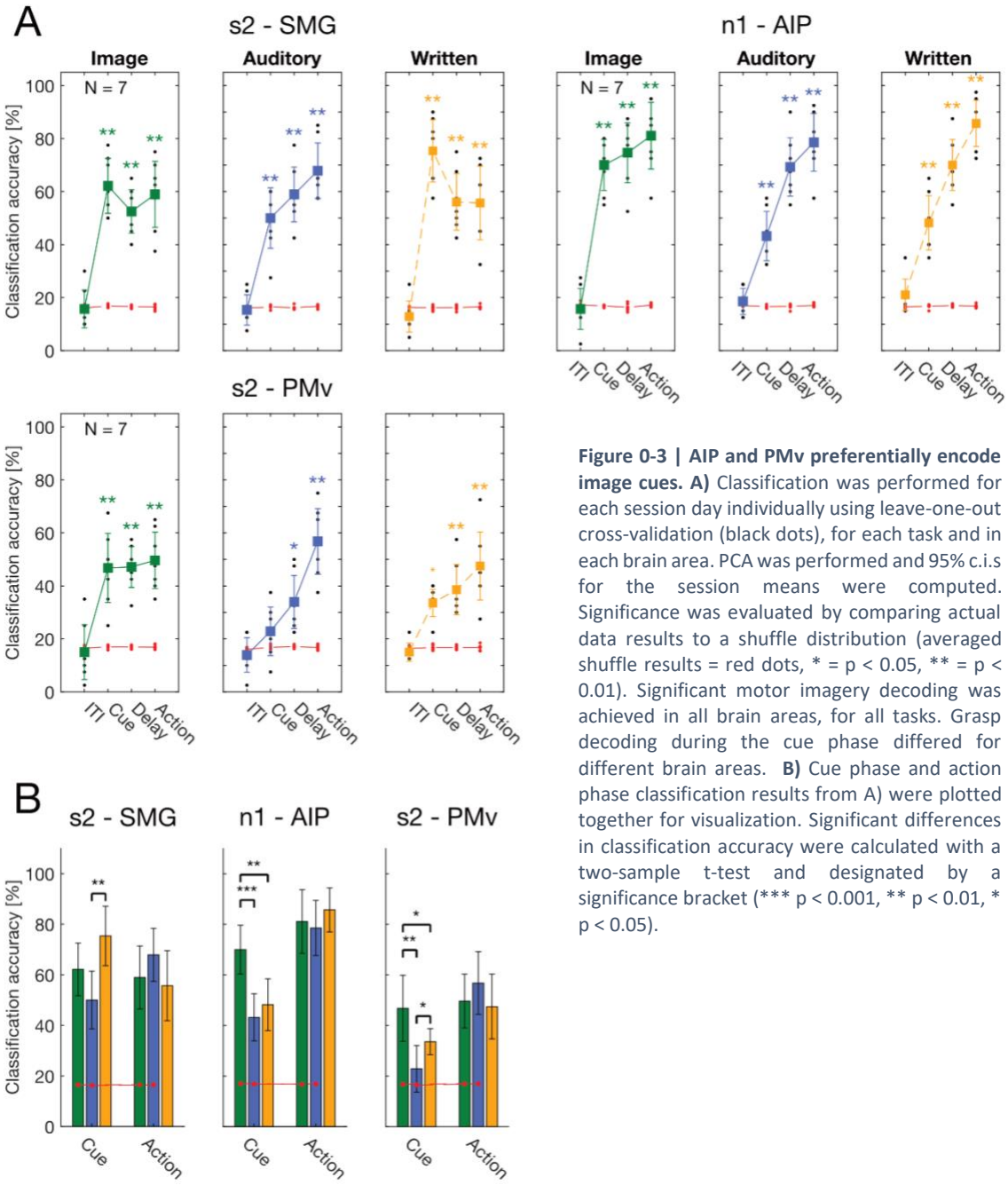

Figure 0-3 | AIP and PMv preferentially encode image cues. A) Classification was performed for each session day individually using leave-one-out cross-validation (black dots), for each task and in each brain area. PCA was performed and 95% c.i.s for the session means were computed. Significance was evaluated by comparing actual data results to a shuffle distribution (averaged shuffle results = red dots, * = $p < 0.05$, ** = $p < 0.01$). Significant motor imagery decoding was achieved in all brain areas, for all tasks. Grasp decoding during the cue phase differed for different brain areas. **B)** Cue phase and action phase classification results from A) were plotted together for visualization. Significant differences in classification accuracy were calculated with a two-sample t-test and designated by a significance bracket (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$).

Sensory modality specific grasp encoding varies in SMG and AIP

The decoding analysis confirmed that regardless of the employed cue modality, grasp motor-imagery could significantly be decoded in all brain areas. Furthermore, grasps could significantly be decoded during the cue phase for all three sensory modalities in SMG and AIP, albeit with different

decoding accuracies. To characterize which cognitive process the neuronal populations represent during the cue phase, we assessed the similarity of information content between the image cue, the auditory cue, and the written cue through a cross-task classification analysis. PMv was excluded from the analysis, as it only weakly or not significantly encoded written and auditory sensory modalities respectively.

Cross-task classification consists of training a classification model on data recorded during the cue phase of one of the tasks, e.g., the "Image cue" task, and using that model to decode grasps during the cue phase of all three tasks. **Figure 4-4A** illustrates some potential outcomes of the analysis. If activity during the cue phase is completely sensory modality dependent, a model trained on image data might not generalize at all to auditory and written cue data. This outcome could suggest neural activity during the image cue is due to a visual process (**Figure 4-4A,** possible outcome 1). On the other hand, if a model trained on the cue phase were to generalize with no significant difference in classification accuracy between the three sensory modalities, it would suggest sensory modality independent activity. Semantic, memory, or planning processes may generate such an outcome (**Figure 4-4A,** possible outcome 2). Finally, if the model would partly generalize to auditory and written cue data, but with lower classification accuracy, that may indicate a mix of sensory and generalized activity (**Figure 4-4A,** possible outcome 3). The analysis was performed by varying the task used for training the model. As a control, the analysis was performed with ITI phase data and significance was assessed by comparing data to a shuffled distribution. Similarity in information content during the cue phase was assessed through a two-sample *t*-test ($p < 0.05$).

We found higher sensory cue-specific activity in SMG than in AIP (**Figure 4-4B**). Indeed, models trained on neural data recorded during one of the cue modalities had significantly lower classification accuracies when tested on data recorded during the other cue modalities in SMG. Interestingly, models trained on auditory cue generalized better to written cue data, and vice versa, suggesting a shared cognitive process between hearing the name of a grasp and silently reading the name of the grasp. In AIP, cue-dependent neural encoding was only observed for the image cue condition. When training a model on auditory cue or written cue data, the model decoded grasps in the other cue modalities equally well. These findings suggest AIP represents a shared neural cognitive process during grasp preparation.

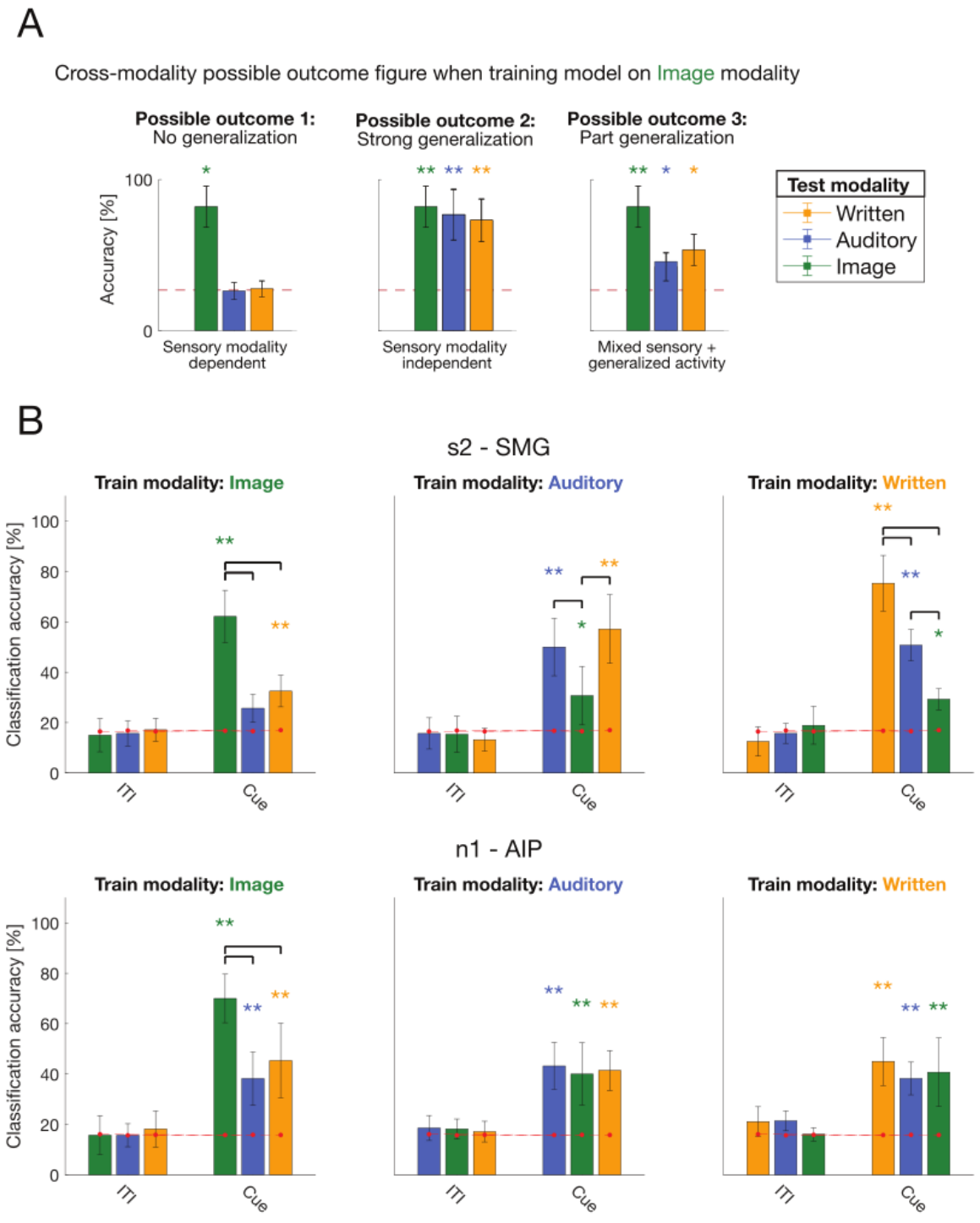

Figure 0-4 | Shared neural representations between image, auditory, and written cue processes. A) Schematic illustrating expected cross-task classifier behavior if neural cue phase activity is 1) sensory modality dependent, 2) sensory modality independent, or 3) mixed sensory and generalized activity. **B)** Performance of decoders trained on data of each cue modality (Image, Auditory, Written), for each brain area in ITI and Cue phase. Significant classification

result compared to shuffled chance distribution are denoted in ** or * (** $p < 0.01$, * $p < 0.05$). Significant differences in classification accuracy were calculated with a two-sample t-test and designated by a classification bracket if $p < 0.05$.

We evaluated the generalizability of grasp motor imagery cued with different sensory modalities by computing a neuron dropping curve combined with cross-task classification. This method allowed investigation of how motor representation for grasp encoding evolves when adding units to the pool of predictors. Results were averaged over 8-folds and bootstrapped confidence intervals (c.i.s.) of the mean were computed over 100 repetitions (**Figure 4-5)**. We found high generalizability of the different models, suggesting grasp motor imagery can be cued successfully with a variety of sensory modalities. Furthermore, classification accuracies in SMG and PMv were consistent with findings from an earlier study (**Figure 3-4**, Chapter 3), indicating robust grasp motor imagery representation.

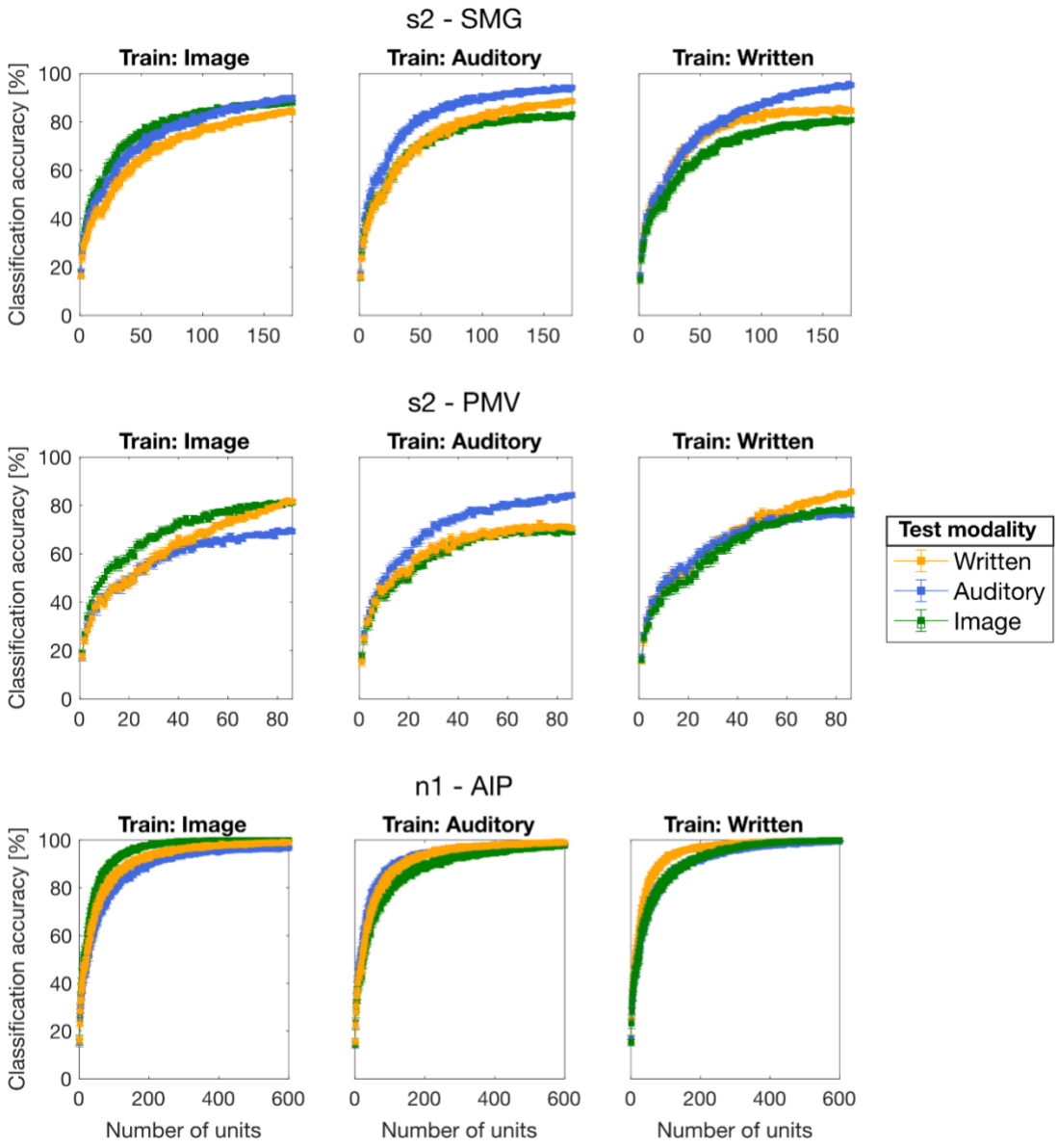

Figure 0-5 | Grasp motor imagery is stable across cue modalities. Neuron dropping curves were performed in SMG, PMv and AIP over 100 repetitions of eight-fold cross validation. The first 20 PCs were used as features. The model was trained on action phase data from the Image cue task (e.g., Train: Image) and tested on action phase data of all tasks. The analysis was repeated by training the model on the Auditory cue task (Train: Auditory) and the Written cue task. The mean classification accuracy with bootstrapped 95% c.i.s. of the mean are plotted.

Discussion

In this work, we showed that participants with tetraplegia can robustly generate cue-modality independent grasp motor imagery. Different higher-level brain regions of the cortical grasp circuit

showed preferential representation of different cue modalities. While PMv and AIP encoded images of grasps significantly higher than audio or written word grasp representations, SMG exhibited sensory-specific encoding for all cue modalities. In SMG, auditory and written shared stronger neural representation, potentially indicating linguistic processes.

Grasp motor imagery is cue-independent

Intracortical recordings from the cortical grasp circuit exhibit planning and execution activity during motor imagery of different grasp shapes using visual cues (image of grasp). However, these higher-level brain areas are also involved in other tasks, such as visual word recognition and phonological processing. We aimed to understand if these different sensory and behavioral paradigms could potentially modulate grasp preparatory and motor activity, by cueing tetraplegic participants to perform motor imagery using three different sensory cues, an image cue, an auditory cue, and a written cue. First, by computing the percentage of tuned units to grasp motor imagery over time (**Figure 4-2**), we found that temporal dynamics during motor imagery were preserved in each task. Through classification analysis, we demonstrated that evoked grasp motor imagery not only was similarly strong in each task (**Figure 4-3B**), but also preserved the same neural code (**Figure 4-5**). These results confirm visual images of grasps are not necessary to generate significant grasp motor imagery in areas of the cortical grasp circuit.

Other studies have evaluated different cues for motor applications, suggesting cue-independent motor imagery for other motor processes as well (T. Aflalo et al. 2020; Guan et al. 2022). In (T. Aflalo et al. 2020), motor imagery of finger flexion was cued with abstract images as a control, and resulted in similar temporal dynamics to motor imagery cued with action verbs (Figure S10). In (Guan et al. 2022), motor imagery of finger presses were cued with interleaved text and spatial cues. However, to our knowledge, the evoked motor imagery using different cues was not directly compared as in our study.

Unique sensory representation in brain areas of the cortical grasp circuit

SMG, AIP, and PMv are part of the cortical grasp circuit. Our finding suggest they may process different cognitive processes during grasp preparation. AIP and PMv preferentially encode image cues depicting a grasp and an object, demonstrated by a higher percentage of tuned units (**Figure**

4-2) and significantly higher classification accuracies (**Figure 4-3B**). These results are consistent with non-human primate studies, which have shown AIP and PMv (putative NHP areas F5) strongly respond to visual presentations of objects and to their size, shape, and orientation (Schaffelhofer and Scherberger 2016; A. Murata et al. 2000; Taira 1998; Baumann, Fluet, and Scherberger 2009). Additionally, Klaes et al. found in human AIP that when grasps were simultaneously cued with non-congruent images and auditory stimuli, only the image cue was decodable (Klaes et al. 2015). These studies further highlight AIP's preferred role in the processing of visual cues.

While significant classification of the auditory and the written cue were found in AIP (**Figure 4-3A**), that activity emerged later compared to the image cue (**Figure 4-2**) and did not appear specific to the sensory cues. Indeed, when training classifier models on these datasets, strong shared neural representations were found (**Figure 4-4**, Train modality: Auditory, Train modality: Written), suggesting a shared cognitive process was represented in all three cue modalities. These results could be consistent with findings in (Schaffelhofer and Scherberger 2016), showing AIP maintained coding for visual object properties during movement execution in the dark, suggesting involvement in working memory. (T. Aflalo et al. 2020) found that passive viewing of abstract symbols did not evoke neural selectivity in PPC, while passive viewing of action verbs did. Further investigation is required to establish if passive viewing of the different cue modalities would evoke neural activity during the cue phase and may elucidate if activity is intrinsic, or task dependent.

SMG neurons showed strongest ability in differentiate written words, demonstrated both by the highest percentage of tuned units (**Figure 4-2**) and significantly higher classification accuracy compared to the other cues (**Figure 4-3B**). However, SMG exhibited cue-specific encoding for the image cue and auditory cue as well. Classification accuracies during auditory and image cue phase were significant (**Figure 4-3A**), and cross-task classification was significantly higher for the cue modality the decoder was trained on compared to other cue modalities (**Figure 4-4B**).These results suggest neurons in SMG can integrate a variety of different sensory modalities. In the literature, we find different possible explanations for this observation. Firstly, studies using functional magnetic imaging (fMRI) posit that SMG functions as a sort of hub in tool use, integrating visual features of the tool, as well as the semantic knowledge of how to use it (McDowell et al. 2018; Johnson-Frey 2004b; Osiurak and Badets 2016; Vingerhoets 2014; Garcea and Buxbaum 2019). These

characteristics may require processing of different sensory modalities during grasping. Secondly, a different body of work from the imaging and TMS literature describes SMG involvement in linguistic processes. Repeatedly, SMG activation or disruption has been observed in tasks involving written word recognition, phonetic, and verbal memory processing (Stoeckel et al. 2009; Sliwinska et al. 2012a; Oberhuber et al. 2016; Deschamps, Baum, and Gracco 2014; Seghier et al. 2004), suggesting SMG's involvement in language. In our work, we observed stronger shared neural representations between auditory and written word grasp representation (**Figure 4-3B**, Train modality: Auditory, Train modality: Written), than for a grasp images. These processes involve silent reading of the cue and auditory understanding of the cue, which require access to language representation, and therefore may suggest SMG processes linguistic aspects. Directly studying SMG's involvement in speech may allow to answer this question.

Impact for BMI applications

In this study, we found that grasp motor imagery was well decodable from SMG, PMV, and AIP. Classification accuracies in SMG and PMv were consistent with findings from an earlier study (Figure 4-5, **Figure 3-4**, Chapter 3). As data from this study were obtained 6 – 10 months later than from the first study, these findings indicate robust and long-term grasp motor imagery representation for BMI applications.

The ability to cue BMIs with a variety of different sensory modalities without affecting decoding abilities can have several advantages. It allows tailoring cues to the individual user's preferences or abilities and may enhance their engagement or satisfaction with the system. For instance, auditory cues do not require the presence of a screen, and can be presented through headphones or speakers, making cue presentation quicker, more private, and more efficient. Furthermore, understanding which cognitive processes are represented in the location of BMI implants may inform the development of novel BMI applications.

Conclusion:

In this work, we studied the effect of the sensory cue modality on grasp motor imagery in different areas of the cortical grasp circuit. Our findings show motor imagery can be decoded with high accuracy regardless of the employed cue modality. Moreover, we show neurons in PMv and AIP preferentially encoded visual images. AIP neurons furthermore showed cue-independent

modulation, suggesting representation of a common cognitive process such as working memory (Schaffelhofer and Scherberger 2016). In SMG, neurons showed cue-specific activity for image, auditory, and written grasp representation. Additionally, shared neural generalization between auditory and written cue processes suggest SMG's involvement in linguistic processes (Oberhuber et al. 2016; Deschamps, Baum, and Gracco 2014). These findings provide new insights into the neural mechanisms underlying motor imagery and other cognitive processes. To conclude, these results suggest BMI users could tailor cues to individual users' preferences and could help guide new applications for BMI devices.

Acknowledgements

We wish to thank L. Bashford, H. Jo, and I. Rosenthal for helpful discussions and data collection. We wish to thank T. Aflalo, K. Kadlec, and C. Guan for data collection with n1. We wish to thank our study participants s2 and n1 for their dedication to the study which made this work possible. This research was supported by the NIH National Institute of Neurological Disorders and Stroke Grant U01: U01NS098975 (S.K.W., S.K., D.B., K.P., C.L. and R.A.A.) and by the T&C Chen Brain-Machine Interface Center (S.K.W., D.B., R.A.A.).

Author contributions

S.K.W, S.K., and R.A.A. designed the study. S.K.W. and S.K. developed the experimental tasks. S.K.W., S.K., and D.B. analyzed the results. S.K.W., S.K., D.B., and R.A.A. interpreted the results and wrote the paper. K.P. coordinated regulatory requirements of clinical trials. C.L. and B.L. performed the surgery to implant the recording arrays.

Chapter 5:

Grasp and speech motor imagery representation in the supramarginal gyrus

The following chapter's contents are taken and adapted from Wandelt et al. 2022, with modifications done to fit the dissertation format. Work of this paper comparing grasp and speech processes is covered in this chapter. For work related to grasp signals, see Chapter 3.

Wandelt, S. K., Kellis, S., Bjånes, D. A., Pejsa, K., Lee, B., Liu, C., & Andersen, R. A. (2022). Decoding grasp and speech signals from the cortical grasp circuit in a tetraplegic human. Neuron, 110(11), 1777-1787

Abstract

Brain-machine interfaces (BMIs) have the potential to increase the independence of people living with paralysis due to spinal cord injury, stroke, or neurological diseases. BMIs located in the human cortical grasp circuit have shown promising results for decoding intended arm and grasps movement. Recently, studies have shown that certain regions in the cortical grasp circuit also encodes useful signals for BMI speech applications. In this study, we investigated three brain areas in the ventral premotor cortex (PMv), the posterior parietal cortex (PPC) and the somatosensory cortex (S1) for grasp motor imagery and vocalized speech signals on the single neuron level. A tetraplegic participant performed grasp motor imagery and spoke words related to grasping and colors while we recorded neural activity in the supramarginal gyrus (SMG), PMv, and the arm areas of S1. We found that all brain areas showed modulation for grasp processes, but only SMG significantly encoded spoken grasp and spoken color words. We found evidence that SMG engages different motor plans for speech and grasp processes by analyzing which neuronal populations were actively involved in both processes, and through cross classification. These findings suggesting a BMI could be trained to represent both grasp and speech processes at the same time. Overall, our study provides insights into the types of neural signals that can be recorded in the cortical grasp circuit, and suggest that neural signals in high-level brain areas in the human cortex can be exploited for grasp and speech related BMI applications.

Introduction

Brain-Machine Interfaces (BMI's) represent a novel technological pathway to bypass functional losses due to paralysis caused by spinal cord injuries, strokes or diseases. People with tetraplegia have consistently rated recovery of hand and arm function as the highest priority for increasing their quality of life (Anderson 2004), (Snoek et al. 2004). Similarly, patients suffering from certain neurological disorders, such as amyotrophic lateral sclerosis (ALS), consider loss of speech among the worst aspects of their disease (Hecht et al. 2002).

Extensive previous work has shown that the human cortical grasp circuit is an excellent target for intracortical BMI applications, allowing to read out reaching and grasping signals (Tyson Aflalo et al. 2015; Collinger et al. 2013; Klaes et al. 2015; Wandelt et al. 2022a; Guan et al. 2022; F. R. Willett et al. 2021). However, recent work has shown that a variety of other sensorimotor signals can be decodable from neurons recorded in this pathway. (Stavisky et al. 2019) and (Wilson et al. 2020) demonstrated speech decoding from the "hand knob" area in M1, an important step towards speech-BMIs for people suffering from neurological disorders or strokes. Evidence for language processing has also been documented in PPC. (T. Aflalo et al. 2020) found PPC activation for reading action words and (Zhang et al. 2017) for spoken words. In SMG, transcranial magnetic stimulation (TMS) and fMRI studies have extensively documented its involvement in language processing (Stoeckel et al. 2009), (Sliwinska et al. 2012a), (Oberhuber et al. 2016) and verbal working memory (Deschamps, Baum, and Gracco 2014), suggesting potential involvement in speech production. Furthermore, we found neurons in SMG share neural representation of auditory and written sensory cues, further suggesting involvement in language processing (Chapter 4). Using ECoG, (J. G. Makin, Moses, and Chang 2020) showed that electrodes over SMG contributed to speech decoding. However, to our knowledge, speech decoding has not previously been demonstrated from spiking activity in SMG alone.

In this work, we aimed to further characterize how different brain areas of the cortical grasp circuit represent grasp and speech motor processes. A tetraplegic participant performed motor imagery or performed verbal speech of words related to grasping and to colors, while neurophysiological responses were captured from three implant sites using recording microelectrode arrays, the

supramarginal gyrus (SMG), the ventral premotor cortex (PMv) and the primary sensory cortex (S1). We hypothesized SMG would modulate during speech processes, due to its involvement in visual word recognition (Chapter 4) and phonological processing (Oberhuber et al. 2016).

Methods

For "Data and code availability", "Experimental model and subject details", "Method details", "Data collection" and "Quantification and statistical analysis" paragraphs, see Chapter 3, as experiment were published together. This method section only contains novel methods that have not been described in Chapter 3.

Experimental task

We implemented three tasks that evaluated grasp and speech processes in areas of the human cortical grasp circuit. The "Motor Imagery" task cued five different grasps with visual images taken from the "Human Grasping Database" (Feix et al. 2016) to examine the neural activity related to imagined grasps in SMG , PMv and S1. The grasps were selected to cover a range of different hand configurations and were labeled "Lateral", "WritingTripod", "MediumWrap", "PalmarPinch", and "Sphere3Finger" **(Figure 5-1A)**. The "Spoken Grasps" tasks instructed the participant to vocalize the name of the different grasps, while the "Spoken Colors" task instructed the participant to vocalize the name of different colors.

Motor imagery task

Each trial consisted of four phases, referred to in this paper as ITI, cue, delay, and action **(Figure 5-1B).** The trial began with a brief inter-trial interval (2 sec), followed by a visual cue of one of the five specific grasps (4 sec). Then, after a delay period (gray circle onscreen; 2 sec), the participant was instructed to imagine performing the cued grasp with his right (contralateral) hand (Go trials; green circle on screen; 4 sec). Three datasets had a longer action phase. For these, only data from the first four seconds of the action phase were included in the analysis.

Spoken grasps task

A speaking variation of the task was constructed with the same task design outline above, but instead of performing motor imagery during the action phase, the participant was instructed to vocalize once the name of the grasp.

Spoken colors task

Another variation of this speaking task used five squares of different colors instead of five grasps, and the participant was instructed to vocalize once the color during the action phase **(Figure 5-1A,B)**.

On each of five total session days, a "Motor Imagery task", a "Spoken Grasps task" and a "Spoken Colors task" was performed, to allow comparisons between tasks.

The participant was situated 1 m in front of a LED screen (1190 mm screen diagonal), where the task was visualized. The task was implemented using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) extension for MATLAB (MATLAB. (2018). 9.7.0.1190202 (R2019b). Natick, Massachusetts: The MathWorks Inc.).

Quantification and statistical Analysis

Cross-task classification

To evaluate the similarity of neuronal firing in the "Motor Imagery", the "Spoken Grasps" and the "Spoken Colors" tasks, cross-task classification was performed. This method consisted of training a classifier on the averaged neuronal firing rates recorded during one of the tasks (e.g., "Motor Imagery"), and evaluating it on the neuronal firing rates of all three tasks. For "Spoken Colors", data was only averaged over the first 2s of the cue phase, as neuronal activity for this condition was shorter than for the other tasks. A LDA with PCA and leave-one-out cross validation was performed for each individual phase (see Methods section "Classification").

Combined classification

To evaluate if a classification model can be trained to represent motor imagery and speech processes at the same time, multiunit activity from 96 channels was combined over five session days, resulting in 40 trials per condition. Five motor imagery, five spoken grasps and five spoken colors were tested, leading to 15 classes in total. A multiclass support vector machine (SVM) algorithm with radial basis function kernel was used to train a model and evaluated with 16-fold cross validation. A confusion matrix was computed showing classification accuracies per condition.

Results

SMG significantly decodes spoken grasps and colors

Grasp and speech representations in SMG, PMv, and S1 were characterized by implementing three tasks that cued a human participant to perform motor imagery of five different grasps, vocalize five grasps or vocalize five colors (**Figure 5-1A**). The tasks contained four phases: an inter-trial interval (ITI), a cue phase, a delay phase, and an action phase, during which the participant performed motor imagery or vocalized the cues word. By comparing each region's evoked activity between motor imagery and speech cognitive processes, we aimed to uncover evidence for language-processing activity at the single unit level. During each session, a "Motor Imagery", a "Spoken Grasps" and a "Spoken Colors" version of the task were run (**Figure 5-1A,B,** see Methods). Importantly, both the "Motor Imagery" and the "Spoken Grasps" task were cued with the same images. This allowed us to investigate if the cue representation of the grasps remained similar, even if different motor outputs (grasping vs. speaking) were planned.

Classification results during the action phase corroborate SMG's involvement during language processing (**Figure 5-1C**) (Oberhuber et al., 2016), (Deschamps, Baum, and Gracco 2014), (Stoeckel et al. 2009) (Chapter 4). In the motor imagery task, neurons in SMG, PMv, and S1 were able to significantly decoded grasps. However, only SMG showed significant classification results during vocalization of grasp names and colors. A confusion matrix demonstrated the different grasp motor types, spoken grasp names, and colors were well represented within SMG's neuronal population (**Figure 5-1D**).

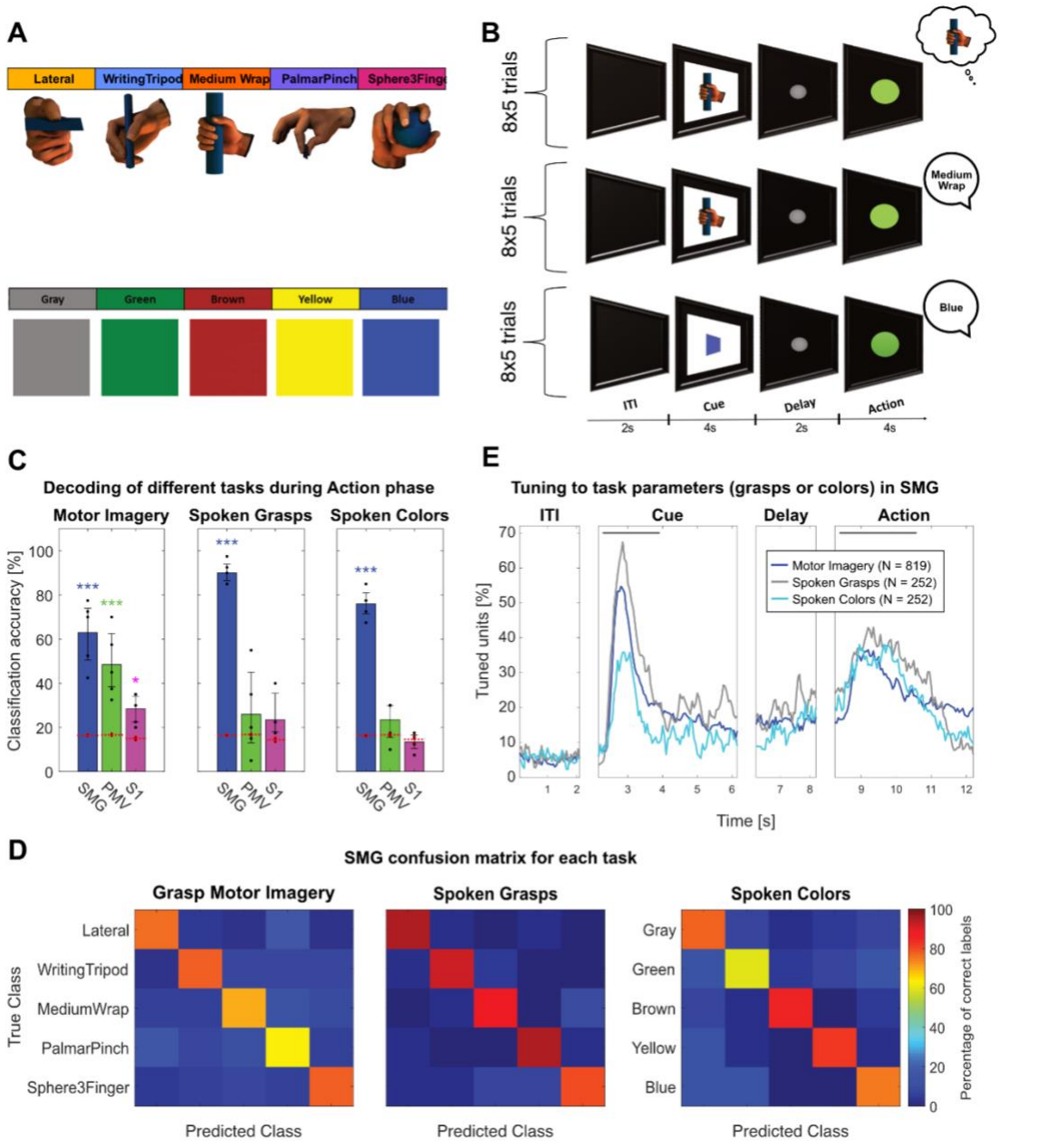

Figure 0-1 | SMG encodes speech. Three tasks were designed studying grasp and speech processes, a "Motor Imagery" task, a "Spoken Grasps" task, and a "Spoken Colors" task. **A)** Grasps images were used to cue the "Motor Imagery" and "Spoken Grasps" tasks. Colored squares were used to cue the "Spoken Colors" task. **B)** The tasks were composed of an inter-trial interval (ITI), a cue phase displaying the image of one of the grasp or colored squares, a delay phase and an action phase. During the action phase, the participant was instructed to perform grasp motor imagery in the "Motor Imagery" task, say out loud the name of the cued grasp during the "Spoken Grasps" task or vocalize the name of the color ("Spoken Colors" task). **C)** Classification was performed for each session day individually using leave-one-out cross-validation (black dots) for "Motor Imagery", "Spoken Grasps" and "Spoken Colors" task. 95% c.i. for the session mean was computed. Results during the action phase are shown. Significance was computed by comparing actual data results to a shuffle distribution (averaged shuffle results = red dots, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$). SMG, PMv, and S1 showed significant classification results when motor imagery was performed. Only SMG showed significant

classification results during spoken grasps and spoken colors. **D)** SMG confusion matrix for each of the tasks. Results were averaged over all session days. **F)** Percentage of tuned units to grasps or colors in 50ms time bins in SMG for each task. The gray lines represent cue and action analysis windows for Figure 6A and B.

To assess selectivity of SMG neurons to the different task parameters, tuning in 50ms bins was computed using linear regression analysis for each task **(Figure 5-1E).** As the exact identification of motor imagery onset is not feasible (having no observable behavioral output), neural data was not aligned to spoken word onset during the speech task, keeping analysis as similar as possible. The population analysis revealed similar temporal dynamics during the cue phase for the "Motor Imagery" and "Spoken Grasps" tasks. This result was expected; both conditions employed the same grasp cue. However, responses for the "Spoken Colors" cues were shorter in time and of lower amplitude, even though they were presented for the same duration as the grasp cues on the screen. During the action phase, temporal dynamics between motor imagery and spoken words were comparable, possibly indicating similar underlying cognitive processes.

We evaluated this hypothesized similarity between motor imagery and speech production by investigating if similar neuronal populations were active in both tasks, and by performing cross-task classification. Tuning analysis allowed us to investigate if similar neural populations were engaged in both tasks. Cross-task classification additionally allowed us to probe if the neural code remained the same.

Using a Venn diagram, we assessed the overlap of tuned units to their respective task parameters during the cue phase **(Figure 5-2A**), and during the action phase **(Figure 5-2B)**. During the cue phase, a high overlap of units active both in the "Motor Imagery" and "Spoken Grasps" task (29%) indicated similar neuronal populations were engaged. This suggests the neuronal representation of the grasp images stayed consistent regardless of the planned output modality **(Figure 5-2A).** Analysis of the action phase activity allowed us to probe if the output modality (speech vs. motor imagery) or the semantic content (grasps vs. colors) were represented by more similar neural populations **(Figure 5-2B).** We found that a notably higher number of units were active during both speech conditions (31%), than between grasp speech and motor imagery condition (8.2%). These results suggest that the output modality drives similarity among neuronal subsets more than semantic content. However, do these findings generalize to classification?

Cross-task classification involved training a classification model on the neuronal firing rate observed in one task, and evaluating the model on all three tasks, performed separately for each phase (see methods). We averaged firing rates over the entire phase duration except during the cue phase for "Spoken Colors". Only the first 2s were analyzed, as neuronal activity for this condition was shorter than for the other tasks **(**Figure 5-1E). During the cue phase, decoding of grasps nicely generalized between the "Motor Imagery" and the "Spoken Grasps" task (**Figure 5-2C**, Train: Motor Imagery; Train: Spoken Grasps). This effect weakened during the delay phase, potentially indicating the formation of separate motor plans for speech and motor imagery. During the action phase, generalization between grasp motor imagery and grasp speech was weak or absent, even if the semantic content was identical. No generalization between the "Spoken Colors" and "Spoken Grasps" tasks occurred, demonstrating that even if similar neuronal populations are active during speech (**Figure 5-2B**), the population activation patterns are different for each spoken word.

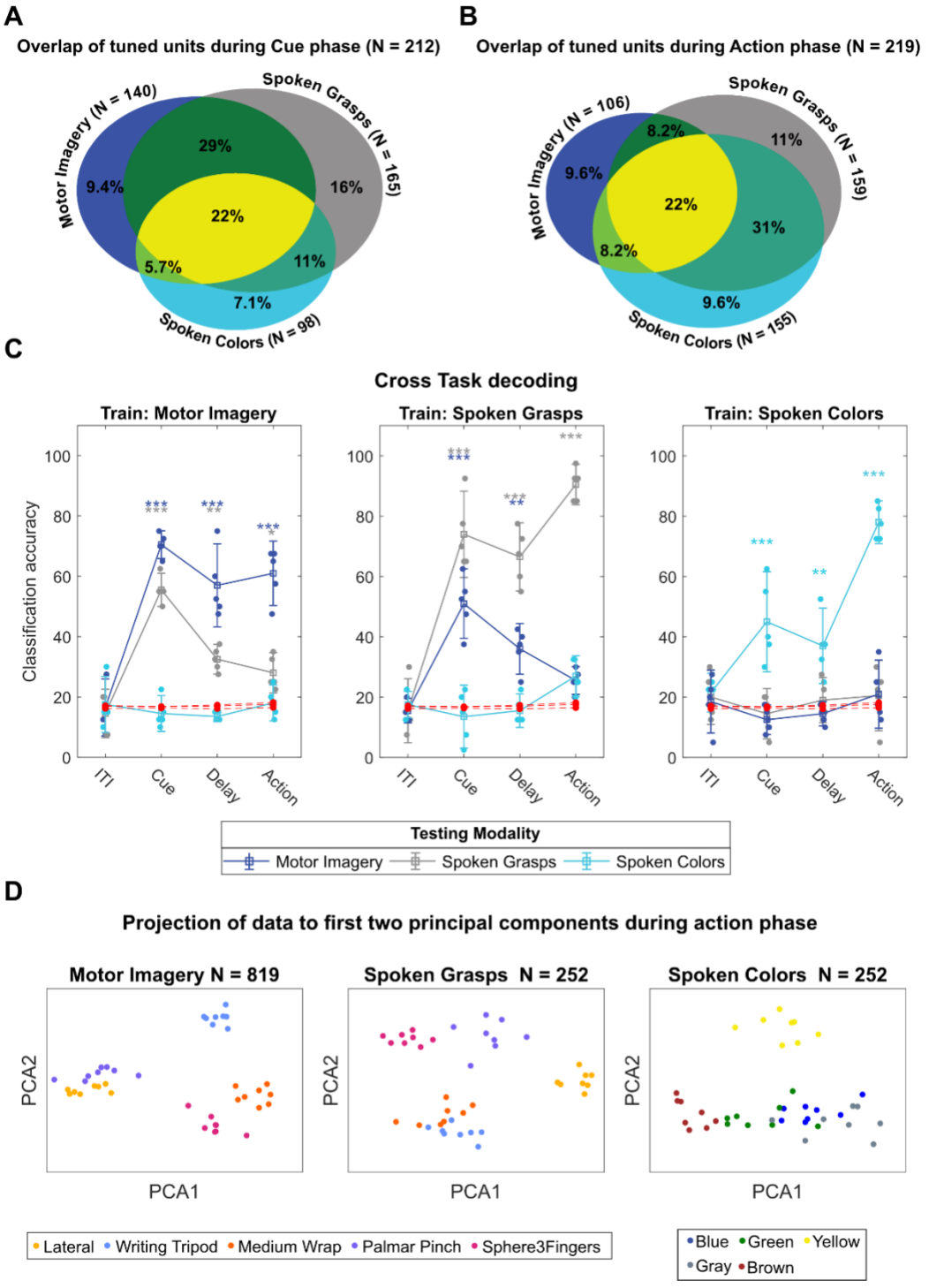

Figure 0-2 | Unique neural representations of motor imagery and speech production in SMG. A) Overlap of units tuned to their respective task parameters during the cue phase between "Motor Imagery", "Spoken Grasps" and "Spoken Colors" task. The overlap of units active for both the motor imagery and the spoken grasps task was notably

higher (29%) compared to motor imagery and spoken colors (5.7%) and spoken colors and spoken grasps (11%). **B)** Same as A but during the action phase. Here, the overlap of units tuned both to spoken grasps and spoken colors is highest (31%). **C)** Cross-task classification was performed by training a classification model on one task (e.g., Motor Imagery) and evaluating it on all three tasks, for each phase separately. Confidence intervals and significance were computed as described in Figure 5C). During the cue phase, generalization between tasks using the same image cue ("Motor Imagery" and "Spoken Grasps") was observed. During the action phase, weak (*) or no generalization was observed. **D)** PCA of the z-scored action phase data was computed for each task. Data was projected onto the first two principal components. Clustering of task parameters (grasps for "Motor Imagery" and "Spoken Grasps", colors for "Spoken Colors") was not consistent between tasks, demonstrating that speech and motor imagery are represented differently in SMG.

We compared encoded feature spaces of the neuronal population during the action phase of the three tasks. The principal components of each z-scored action phase dataset were computed and projected onto the first two principal components. Results are depicted in **Figure 5-2D**. During motor imagery, "Lateral" and "PalmarPinch" occupied neighboring areas in the feature space, suggesting they are more similarly represented in the neural data. During speech conditions, we observed different clustering. Here, "MediumWrap" and "WritingTripod" are represented closer to each other, as well as "PalmarPinch" and "Sphere3Finger". The word with fewest syllables ("Lateral") is represented the furthest away in the neural space. For Spoken Colors, most words were intermixed.

To assess similarities between the cue and the action phase during spoken tasks, we performed a neuron dropping analysis combined with cross-phase classification. Results were similar to the motor imagery results when training on the action phase (See Chapter 3, **Figure 3-4**). When training on the cue phase, seemingly less generalization occurred between the cue and action phase during the "Spoken Grasps" task than during the "Motor Imagery" task (**Figure 5-3A**). For the "Spoken Colors" task, the neural code was similar during the cue and action phases when training on the cue phase; however, it reached a lower average maximum classification accuracy compared to motor imagery and grasp speech (an accuracy of 71% when training on the cue phase of spoken colors vs. 93% for the cue phase of motor imagery, and 95% for the cue phase of spoken grasps) (**Figure 5-3B**).

Finally, we evaluated how well a classification model can distinguish grasp and speech processes simultaneously. Multiunit data from five session days was combined, resulting in 40 trials for each of the 15 conditions (5 grasp motor imagery, 5 spoken grasps, 5 spoken colors). A SVM algorithm

was used to train the model and evaluated with 16-fold cross validation and a confusion matrix was calculated. We found all conditions were well separable from each other, suggesting a classification model may be trained to represent both grasp and speech processes at the same time.

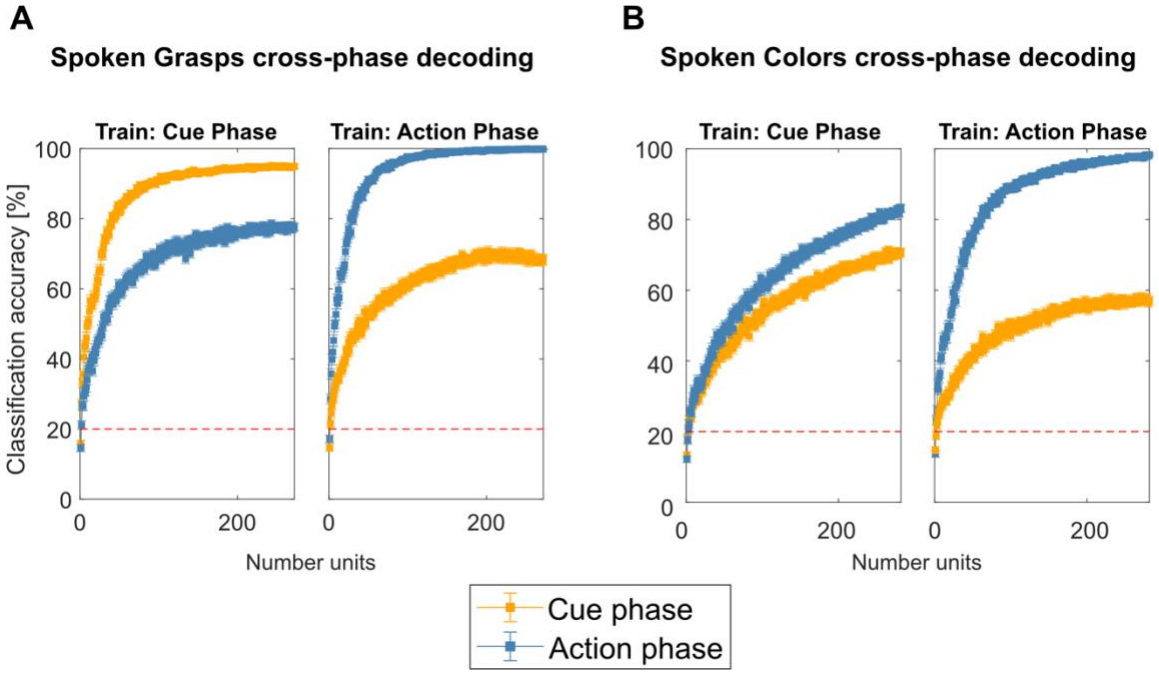

Figure 0-3 | SMG shows less generalizability between the cue and action phase during spoken words than motor imagery. A) A neuron dropping curve analysis was performed on SMG activity using data from the "Spoken Grasps" task over 100 repetitions of eight-fold cross validation. To avoid overfitting, the first 20 PCs were used as features for classification. The analysis was performed once by training the model on the cue phase, and applying it to both cue and action phases (Train: Cue phase), and by training it on the action phase and applying it on both cue and action phase (Train: Action phase). The mean classification accuracy with bootstrapped 95% c.i. was plotted. Less generalization occurred between the cue and action phase during the "Spoken Grasps" task than during "Motor Imagery" task (Figure 3) **B)** Same as A), but for "Spoken Colors" task. Training on the cue phase reached a lower average classification accuracy (71%) compared to the "Motor Imagery" (93% - Chapter 3, Figure 3-4) and "Spoken Grasps" (95%) tasks. This suggests color might be a less important feature than grasp images in SMG.

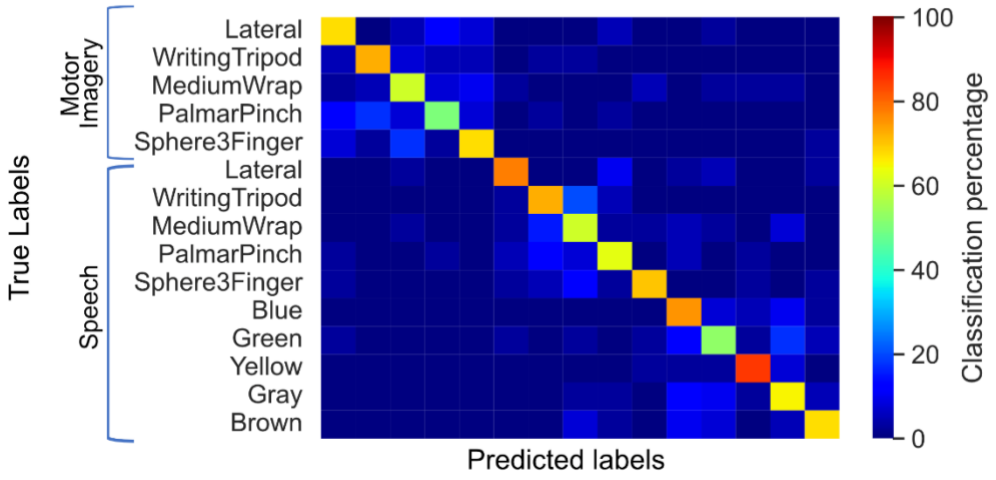

Figure 0-4 | Confusion matrix for combined tasks. For this analysis data from all session days were combined, resulting in 600 trials (200 for grasp motor imagery, 200 for spoken grasp words and 200 for spoken color words). Support vector machine (SMV) classification with radial basis function kernel was performed and evaluated with 16 fold cross-validation.

Discussion

In this work, we showed how speech and grasp motor imagery are represented in the neural populations of SMG, PMv, and S1 of a chronically implanted tetraplegic participant. We found PMv and S1 preferentially represent grasping, while SMG robustly represents spoken grasp and color names. While temporal dynamics for both motor processes were similar, different motor plans were engaged for speaking vs. motor imagery. These characteristics allowed a classifier to represent both processes at the same time, suggesting SMG as a target site for grasp and speech BMI applications.

SMG encodes speech

During speech, SMG and PMv showed vastly different results. Spoken words (both grasp names and colors) were decodable equally or better than only motor imagery of grasps in SMG. In contrast, PMv and S1 showed neither significant classification of spoken grasp names nor of spoken colors (**Figure 5-1C**).

Tuning to task parameters in 50ms time bins allowed us to observe how neuronal activity changed over trial duration for all three tasks. Interestingly, tuning to different color squares was lower in

amplitude than to tuning of grasp images (**Figure 5-1E**), indicating that color might be a less important feature than the shape and size of different grasps in SMG.

The motor imagery and speech tasks showed strong similarities in tuning during the action phase (**Figure 5-1E**). These results could indicate that SMG processes semantics, regardless of the performed task. To answer this question, we performed two different analyses: cross-task tuning, and cross-task classification.

Distinct SMGs encoding for grasp and speech processes

To observe similarities between the neuronal subpopulations involved in the task, cross-task tuning was computed. During the cue phase (**Figure 5-2A**), a high overlap of units tuned during both "Motor Imagery" and "Spoken Grasps" tasks suggests the representation of the cue did not change based on the engaged motor plan. However, during the action phase, (**Figure 5-2B**) this relationship disappeared. We observed a higher overlap of neuronal populations tuned for the output modality (speaking of colors and grasps) than for semantic content (grasp motor imagery and speaking of grasps). At first glance, these results suggest that a similar neuronal population could be active for "Motor Imagery" and "Spoken Grasps" during the cue phase, and for "Spoken Grasps" and "Spoken Colors" during the action phase. However, this evidence does not confirm the neuronal code was identical. Cross-task classification allowed us to answer this question **(Figure 5-2C)**.

During the cue phase, the model generalized nicely between the grasp motor imagery and spoken grasps tasks, confirming the neural code of the grasp image cue remained similar. This effect decreased during the delay phase and became weak or absent during the action phase. These results indicate that SMG engaged different motor plans when motor imagery or speech was performed, even if the meaning of the word remained the same.

During the action phase, none of the models trained during one task generalized to a different task. Even if similar neuronal populations were active during speech **(Figure 5-2B)**, a different activation pattern was observed for each word. Furthermore, accurate classification of color words confirmed SMG's role is not confined to only action verbs, even if classification accuracy of spoken colors was lower than that for spoken grasps. Possibly, the novelty of the words affected the amplitude of

neural representation, as color words are more common than the grasp names we employed. However, our participant was well versed in the names of the grasps, having used them repeatedly during other experiments.

To investigate the underlying feature space encoded in SMG's neuronal population during motor imagery and speech, principal components were computed on the z-scored action phase data of each task. By representing the feature space of the action data in a 2D PCA space, similarities between representations of each grasp could be appreciated through several groupings **(Figure 5-2D)**. During motor imagery, "Lateral" and "PalmarPinch" were grouped closely together. Intuitively, this could be due to a similar pinch with the thumb in these grasp shapes compared to the other grasps. However, the square-shaped objects represented in the grasp images also bear certain similarities. If similarity in object size and shape could predict related population activity, then "MediumWrap" and "WritingTripod" should have been represented in close proximity, as they both depict cylinders of different diameter. This was not observed, as "MediumWrap" was closer to "Sphere3Finger" in the PCA space. Furthermore, in an fMRI study, object size was not shown to modulate SMG activity (Perini et al. 2020). Therefore, we hypothesize that hand posture rather than object shape and size are encoded in the neural data during motor imagery in SMG.

During speech, the representation of the grasps in the neural space changed. This result was expected, as cross-task classification did not generalize well between grasp motor imagery and grasp speech **(Figure 5-2C)**.

Implication for BMI applications

Other studies have characterized simultaneous grasp and speech representation in SMG. Using fMRI, (Andric et al. 2013) showed that left SMG is active both during the observation of grasp actions, emblems (gesture that have a symbolic meaning, e.g. "thumbs-up"), and speech matching the meaning of the emblems. However, authors suggested the activity was not necessarily due to the semantic meaning of the stimuli. These results are consistent with ours, demonstrating SMG activity for both processes, but that semantic meaning was not as strongly encoded as the output modality.

Taken together, these findings suggest neurons in SMG have different neural codes for grasp and speech applications, indicating a BMI could be trained for both processes simultaneously without affecting classification accuracy substantially. **Figure 5-4** showed classes were highly decodable, confirming the model learned well how to differentiate between the different functions. A caveat of the analysis is that data was not recorded as intermixed trials, but in consecutive blocks. By randomizing the order of each task block, and combining data from different session days together, we hope the effect of task classification accuracy was negligible. Nonetheless, these results are promising and suggest a single implant site may allow building grasp and speech BMI applications for people suffering from paralysis.

Conclusion

During speech, SMG achieved significant classification performance, in contrast to PMv and S1, which were not able to significantly decode individual spoken words. While temporal dynamics between motor imagery and speech were similar, we observed different motor plans for each output modality. These results are evidence for a larger role of SMG in language processing. Given the flexibility of neural representations within SMG, this brain area may be a candidate implant site for BMI speech and grasping applications.

Acknowledgements

We wish to thank L. Bashford, H. Jo, and I. Rosenthal for helpful discussions and data collection. We wish to thank our study participant FG for his dedication to the study which made this work possible. This research was supported by the NIH National Institute of Neurological Disorders and Stroke Grant U01: U01NS098975 (S.K.W., S.K., D.B., K.P., C.L., and R.A.A.) and by the T&C Chen Brain-machine Interface center (S.K.W., D.B., R.A.A.).

Author contributions

S.K., S.K.W., and R.A.A. designed the study. S.K.W. and S.K. developed the experimental tasks. S.K.W., S.K. and D.B. analyzed the results. S.K.W., S.K., D.B., and R.A.A. interpreted the results and wrote the paper. K.P. coordinated regulatory requirements of clinical trials. C.L. and B.L. performed the surgery to implant the recording arrays.

Chapter 6:

Online internal speech decoding from single neurons in a human participant

The following chapter's contents are taken and adapted from a preprint by Wandelt et al. 2022 b, with modifications done to fit the dissertation format.

Sarah K. Wandelt, David A. Bjånes, Kelsie Pejsa, Brian Lee, Charles Liu, Richard A. Andersen. Online internal speech decoding from single neurons in a human participant. medRxiv 2022.11.02.22281775; doi: https://doi.org/10.1101/2022.11.02.22281775

Abstract

Speech brain-machine interfaces (BMI's) translate brain signals into words or audio outputs, enabling communication for people having lost their speech abilities due to diseases or injury. While important advances in vocalized, attempted, and mimed speech decoding have been achieved, results for internal speech decoding are sparse, and have yet to achieve high functionality. Notably, it is still unclear from which brain areas internal speech can be decoded. In this work, a tetraplegic participant with implanted microelectrode arrays located in the supramarginal gyrus (SMG) and primary somatosensory cortex (S1) performed internal and vocalized speech of six words and two pseudowords. We found robust internal speech decoding from SMG single-neuron activity, achieving up to 91% classification accuracy during an online task (chance level 12.5%). Evidence of shared neural representations between internal speech, word reading, and vocalized speech processes were found. SMG represented words in different languages (English/ Spanish) as well as pseudowords, providing evidence for phonetic encoding. However, neural data in SMG also separated homophones, suggesting word meaning may also modulate activity. Furthermore, our decoder achieved high classification with multiple internal speech strategies (auditory imagination/ visual imagination). Activity in S1 was modulated by vocalized but not internal speech, suggesting no articulator movements of the vocal tract occurred during internal speech production. This works represents the first proof-of-concept for a high-performance internal speech BMI.

Introduction

Speech is one of the most basic forms of human communication, a natural and intuitive way for humans to express their thoughts and desires. Neurological diseases like amyotrophic lateral sclerosis (ALS) and brain lesions can lead to the loss of this ability. In the most severe cases, patients who experience full-body paralysis might be left without any means of communication. Patients with ALS self-report loss of speech as their most serious concern (Hecht et al. 2002). Brain-machine Interfaces (BMIs) are devices offering a promising technological path to bypass neurological impairment by recording neural activity directly from the cortex. BMIs have demonstrated potential to restore independence to tetraplegic participants by reading out movement intentions directly from the brain (Tyson Aflalo et al. 2015; Andersen et al. 2014a; Andersen, Aflalo, and Kellis 2019; Andersen 2019). Similarly, reading out internal (also reported as inner, imagined, or covert) speech signals could allow the restoration of communication to people who have lost it.

Decoding speech signals directly from the brain presents its own unique challenges. While non-invasive recording methods like functional magnetic imaging (fMRI), electroencephalography (EEG), or magnetoencephalography (MEG) (Dash, Ferrari, and Wang 2020b; Dash et al. 2020) are important tools to locate speech and internal speech production, they lack the necessary temporal and spatial resolution, adequate signal-to-noise ratio, or portability for building an online speech BMI (Luo, Rabbani, and Crone 2022; Martin et al. 2018; Rabbani, Milsap, and Crone 2019). Intracortical electrophysiological recordings have higher signal-to-noise ratios, excellent temporal resolution (Nicolas-Alonso and Gomez-Gil 2012), and are a more suitable choice for internal speech decoding device.

Invasive speech decoding has predominantly been attempted with electrocorticography (ECoG) (Rabbani, Milsap, and Crone 2019) or stereo-electroencephalographic (sEEG) depth arrays (Herff, Krusienski, and Kubben 2020), as they allow sampling neural activity from different parts of the brain simultaneously. Impressive results in vocalized and attempted speech decoding and reconstruction have been achieved using these techniques (Angrick et al. 2018; Herff et al. 2019; Kellis et al. 2010; J. G. Makin, Moses, and Chang 2020; Moses et al. 2021). However, vocalized speech has also been decoded from small-scale microelectrode arrays located in the motor cortex

(Stavisky et al. 2019; Wilson et al. 2020) and the supramarginal gyrus (SMG) (Wandelt et al. 2022a), demonstrating vocalized speech BMIs can be built using neural signals from localized regions of cortex.

While important advances in vocalized speech (J. G. Makin, Moses, and Chang 2020), attempted speech (Moses et al. 2021), and mimed speech (Bocquelet et al. 2016; Anumanchipalli, Chartier, and Chang 2019) decoding have been made, highly accurate internal speech decoding has not been achieved. Lack of behavioral output, lower signal-to-noise ratio, and differences in cortical activation have resulted in much lower classification accuracies of internal speech (Angrick et al. 2018; Martin et al. 2018; Luo, Rabbani, and Crone 2022; Proix et al. 2022). In Pei, Barbour, et al., 2011 patients implanted with ECoG grids over frontal, parietal, and temporal regions silently read or vocalized written words from a screen. Researchers significantly decoded vowels (37.5%) and consonants (36.3%) from internal speech (chance level 25%). (Ikeda et al. 2014) decoded three internally spoken vowels using ECoG arrays using frequencies in the beta band, with up to 55.6% accuracy from Broca area (chance level 33%). Using the same recording technology, Martin et al., 2016 investigated the decoding of six words during internal speech. The authors demonstrated an average pair-wise classification accuracy of 58%, reaching 88% for the highest pair (chance level 50%). These studies were so-called open-loop experiments, in which the data was analyzed offline after acquisition. A recent paper demonstrated real-time (closed loop) speech decoding using stereotactic depth electrodes (Angrick et al. 2021). Results were encouraging as internal speech could be detected; however, the reconstructed audio was not discernable and required audible speech to train the decoding model.

While to our knowledge internal speech has not previously been decoded from SMG, evidence for internal speech representation in SMG exists. In a review of 100 fMRI studies, (Cathy J. Price 2010) described SMG activity not only during speech production, but also suggested its involvement in subvocal speech (Langland-Hassan and Vicente 2018; Perrone-Bertolotti et al. 2014). Similarly, an ECoG study identified high-frequency SMG modulation during vocalized and internal speech (Pei, Leuthardt, et al. 2011b). Additionally, fMRI studies demonstrated SMG involvement in phonologic processing; for instance, the participant decided if two words rhyme (Oberhuber et al. 2016). Performing such tasks requires the participant to internally “hear” the

word, indicating potential internal speech representation (Binder 2017). Furthermore, a study performed in people suffering from aphasia found lesions in SMG and its adjacent white matter affected inner speech rhyming tasks (Geva et al., 2011). Recently, (J. G. Makin, Moses, and Chang 2020) showed electrode grids over SMG contributed to vocalized speech decoding. Finally, vocalized grasps and color words were decodable from SMG from the same participant involved in this work (Wandelt et al. 2022a). These studies provide evidence for the possibility of an internal speech decoder from neural activity in SMG.

The relationship between inner speech and vocalized speech is still debated. The general consensus posits similarities between internal and vocalized speech processes (Pei, Leuthardt, et al. 2011b), but the degree of overlap is not well understood (Cooney, Folli, and Coyle 2022; 2022; Martin et al. 2018; Perrone-Bertolotti et al. 2014; Alderson-Day and Fernyhough 2015). Characterizing similarities between vocalized and internal speech could provide evidence that results found with vocalized speech could translate to internal speech. However, such a relationship may not be guaranteed. For instance, some brain areas involved in vocalized speech might be poor candidates for internal speech decoding.

In this work, a participant with tetraplegia performed internal and vocalized speech of six words and two pseudowords, while neurophysiological responses were captured from two implant sites. We investigated representations of various language processes at the single-neuron level using recording microelectrode arrays from the supramarginal gyrus (SMG) located in the posterior parietal cortex (PPC) and the arm region of the primary somatosensory cortex (S1). Words were presented with an auditory or a written cue, and were produced internally as well as orally. We hypothesized SMG and S1 activity would modulate during vocalized speech and SMG activity would modulate during internal speech. Shared representation between internal speech, vocalized speech, auditory comprehension, and word reading processes were investigated.

Methods

Experimental model and subject details

A tetraplegic participant was recruited for an IRB- and FDA-approved clinical trial of a brain-machine interface and he gave informed consent to participate. The participant suffered a spinal cord injury at cervical level C5 two years prior to participating in the study.

Method details

Implants

The targeted areas for implant were the supramarginal gyrus (SMG), and primary somatosensory cortex (S1) and the left ventral premotor cortex (PMv). In this study, SMG and S1 data was considered. For description of localization fMRI tasks and implant locations see (Armenta Salas et al. 2018). In November 2016, the participant underwent surgery to implant one 96-channel multi-electrode array (Neuroport Array, Blackrock Microsystems, Salt Lake City, UT) in SMG and PMv each, and two 7 x 7 sputtered iridium oxide film-tipped microelectrode arrays with 48 channels each in S1. Data were collected between July 2021 and August 2022.

Data collection

Recording began two weeks after surgery and continued one to three times per week. Data for this work were collected between 2021 and 2022. Broadband electrical activity was recorded from the NeuroPort arrays using Neural Signal Processors (Blackrock Microsystems, Salt Lake City, UT). Analog signals were amplified, bandpass filtered (0.3–7500 Hz), and digitized at 30,000 samples/sec. To identify putative action potentials, these broadband data were bandpass filtered (250-5000 Hz), and thresholded at -4.5 the estimated root-mean-square voltage of the noise. For some of the analyses, waveforms captured at these threshold crossings were then spike sorted by manually assigning each observation to a putative single neuron, for others, multiunit activity was considered. On average, 33 sorted SMG units (between 22–56) and 83 sorted S1 units (between 59–96) were included in the analysis. Auditory data was recorded at 30k Hz simultaneously to the neural data. Background noise was reduced post-recording by using the noise reduction function of the program “Audible”.

Experimental Tasks

We implemented different tasks to study language processes in the SMG. The tasks cued six words informed by Martin et al., 2016 (Spoon, Python, Battlefield, Cowboy, Swimming, Telephone) as well as two pseudowords (Bindip, Nifzig). The participant was situated 1m in front of a LED screen (1190 mm screen diagonal), where the task was visualized. The task was implemented using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007) extension for MATLAB.

Auditory cue task

Each trial consisted of six phases, referred to in this paper as ITI, Cue, D1 (delay 1), Internal, D2 (delay 2), and Speech. The trial began with a brief inter-trial interval (2s), followed by a 1.5s long cue phase. During the cue phase, a speaker emitted the sound of one of the eights words (e.g., Python). Word duration varied between 842 and 1130 ms. Then, after a delay period (gray circle on screen; 0.5s), the participant was instructed to internally say the cued word (orange circle on screen; 1.5 sec). After a second delay (gray circle on screen; 0.5s), the participant vocalized the word (green circle on screen, 1.5 seconds).

Written cue task

The task was identical to the Auditory cue task, except words were cued in writing instead of audio. The written word appeared on the screen for 1.5 seconds during the cue phase. The auditory cue was played ~ 250ms later than the written cue appeared on the screen, due to software constraints.

The auditory cue task and written cue task were recorded on 10 different session days.

Control experiments

Three experiments were run to investigate internal strategies and phonetic vs. semantic processing.

Internal strategy task

The task was designed to vary the internal strategy employed by the participant during the internal speech phase. Two internal strategies were tested: a sound imagination and a visual imagination. For the “sound imagination” strategy, the participant was instructed to imagine what

the sound of the word sounded like. For the "visual imagination" strategy, the participant was instructed to perform mental visualization of the written word. We also tested if the cue modality (auditory or written) influenced the internal strategy. A subset of four words were used for this experiment. This led to four different variations of the task:

Task variation	Cue modality	Internal strategy
Auditory - Sound	Auditory	Sound imagination
Written – Visual	Written	Visual imagination
Auditory – Visual	Auditory	Visual imagination
Written – Sound	Written	Sound imagination

English/Spanish task

The task was designed to understand if SMG stronger encodes semantic or phonetic information. For this task, we made use of the participant's bilingual abilities and asked him to translate a subset of the words into their Spanish counterparts. This ensured the semantic meaning of words remained the same, while varying the phonetic content. Those words were used to design the Spanish version of the task. On each session day, an English and a Spanish version of the task was run.

Homophone task

The task was designed to test if word meaning and context were represented in SMG's neural activity. We tested two groups of three homophones (words that have different semantic meanings but same pronunciation): Scent, Sent, Cent and Ware, Wear, Where.

The internal strategy task was run on one session day, the English/Spanish task was run on three session days, and the Homophone task was run on four session days.

Online task

The "Written cue task" was turned into a closed-loop experiment. To obtain training data for the online task, a written cue task was run. Then, a classification model was trained only on the internal speech data of the task (see classification subsection). The closed-loop task was nearly identical to the "Written cue task" but replaced the vocalized speech phase by a feedback phase.

Feedback was provided by showing the decoded word on the screen either in green if correctly classified, or in red if wrongly classified. See Supplementary Video 1 for an example of the participant performing the online task.

Error trials

Trials were the participants accidentally spoke during the internal speech part (3) or said the wrong word during the vocalized speech part (20) were removed from all analysis.

Quantification and statistical analysis

Analyses were performed using MATLAB R2020b and Python, version 3.8.11.

Neural firing rates

Firing rates of sorted units were computed as the number of spikes occurring in 50ms bins, divided by the bin width, and smoothed using a Gaussian filter with kernel width of 50ms to form an estimate of the instantaneous firing rates (spikes/sec).

Linear regression analysis

To identify units exhibiting selective firing rate patterns (or tuning) for each of the eight words, linear regression analysis was performed in two different ways: 1) step by step in 50ms time bins to allow assessing changes in neuronal tuning over the entire trial duration; 2) averaging the firing rate in each task phase to compare tuning between phases. The model returns a fit that estimates the firing rate of a unit based on the following variables:

$$FR = \beta o + \beta 1X1 + \beta 2X2 + \beta 3X3 + \beta 4X4 + \beta 5X5 + \beta 6X6 + \beta 7X7 + + \beta 8X8,$$

where FR corresponds to the firing rate of the unit, βo to the offset term which was the average ITI data of the unit, and β corresponds to the estimated regression coefficients (see (Wandelt et al. 2022)).

In this model, β symbolizes the change of firing rate from baseline for each word. A student's t – test was performed to test the hypothesis of $\beta = 0$. A follow-up analysis was performed to adjust for false discovery rate between the p-values (Benjamini and Hochberg 1995) (Benjamini and Yekutieli 2001). A unit was defined as tuned if the hypothesis could be rejected (adjusted p-value

< 0.05, t-statistic) for at least one word. This definition allowed for tuning of a unit to zero, one, or multiple words during different time points of the trial. Linear regression was performed for each session day individually. A 95% confidence interval was computed by performing the student's t-inverse cumulative distribution function over the 10 sessions.

Classification

Using the neuronal firing rates recorded during the tasks, a classifier was used to evaluate how well the set of words could be differentiated during each phase. Classifiers were trained using averaged firing rates over each task phase. A model for each phase was built using linear discriminant analysis (LDA), assuming an identical covariance matrix for each word, which resulted in best classification accuracies. Principal component analysis (PCA) was applied on the training data and PCs explaining more than 95% of the variance were selected as features, and applied to the testing set. Leave-one-out cross-validation was performed to estimate decoding performance. A 95% confidence interval was computed as described above.

Cross-phase classification

To estimate shared neural representations between different task phases, we performed cross-phase classification. The process consisted in training a classification model (as described above) on one of the task phases (e.g., ITI) and to test it on the ITI, cue, imagined speech, and vocalized speech phases. The method was repeated for each of the 10 sessions individually, and a 95% confidence interval of the mean was computed. Significant differences in classification accuracies between phases decoded with the same model were evaluated using a paired t-test, with alpha = 0.001.

English/Spanish task classification

To evaluate if semantic or phonetic processes are encoded within SMG, data from an English and a Spanish version of the task were merged. Multiunit activity from 96 channels was combined over three session days, resulting in 24 trials per condition. A multiclass support vector machine (SVM) algorithm with radial basis function kernel was used to train a model on one task phase (e.g., ITI), and to test it on the ITI, cue, imagined speech, and speech phases. 10-fold cross validation was performed for model evaluation. A confusion matrix was computed showing classification accuracies per word.

Classification performance significance testing

To assess the significance of classification performance, a null dataset was created by repeating classification 100 times with shuffled labels. Then, different percentile levels of this null distribution were computed and compared to the mean of the actual data. Mean classification performances higher than the 95th percentile were denoted with a * symbol and higher than 99th percentile were denoted with **.

Demixed principal component (dPCA) analysis

dPCA analysis was performed to break down the activity of the neuronal population into individual components, also called marginalizations, and to observe the explained variance contained in the data for each marginalization. This analysis showed the contribution of each signal variable in the observed data. We followed the method and code (https://github.com/machenslab/dPCA) described in (Kobak et al. 2016)and adapted it to our dataset.

Our dataset had three parameters: timing, cue modality, (e.g., auditory or visual), and word (8 different words). As in the original manuscript, data were decomposed into five parts: condition-independent, cue modality-dependent, word-dependent, dependent on the cue modality-word interaction, and noise. Similar to the covariance decomposition done in ANOVA, individual terms were given by a series of averages. Some of the terms were grouped together as they were individually of lesser interest for this analysis. The Timing marginalization combined the time, and cue modality–word–time interaction. The Cue Modality marginalization combined the cue modality, cue modality-time interaction, and cue modality-word interaction. The Word marginalization combined the word and word-time interaction. To avoid overfitting, a regularization term lambda was used. The analysis was performed as described in in (Kobak et al. 2016).

Results

Task design

We characterized neural representations of four different language processes within a population of SMG neurons: auditory comprehension, word reading, internal speech, and vocalized speech production. In this thesis, internal speech refers to engaging a prompted word internally (“inner

monologue"), without correlated motor output, while vocalized speech refers to audibly vocalizing a prompted word.

The task contained six phases: an inter-trial interval (ITI), a cue phase (Cue), a first delay (D1), an internal speech phase (Internal), a second delay (D2), and a vocalized speech (Speech) phase. Words were cued with either an auditory or a written version of the word (**Figure 5-1A**). Six of the words were informed by Martin et al., 2016 (Battlefield, Cowboy, Python, Spoon, Swimming, Telephone). Two pseudowords (Nifzig, Bindip) were added to explore phonetic representation in SMG.

Single neurons modulate firing rate during internal speech in SMG

For each of the four language processes, we observed selective modulation of individual neurons' firing rates (**Figure 5-1B,C)**. In general, firing rates of neurons increased during the active phases (Cue, Internal, Speech) and decreased during rest phases (ITI, D1, D2). A variety of activation patterns were present in the neural population. Example neurons were selected to demonstrate increases in firing rates during internal speech (**Figure 5-1B**, Spoon); neurons were also active during the cue and vocalized speech (**Figure 5-1B**, Figure 5-S1). Regardless of the cue modality (auditory in **Figure 5-1B,C,** written in **Figure 5-1D,E**), internal speech highly modulated individual neuron firing rates.

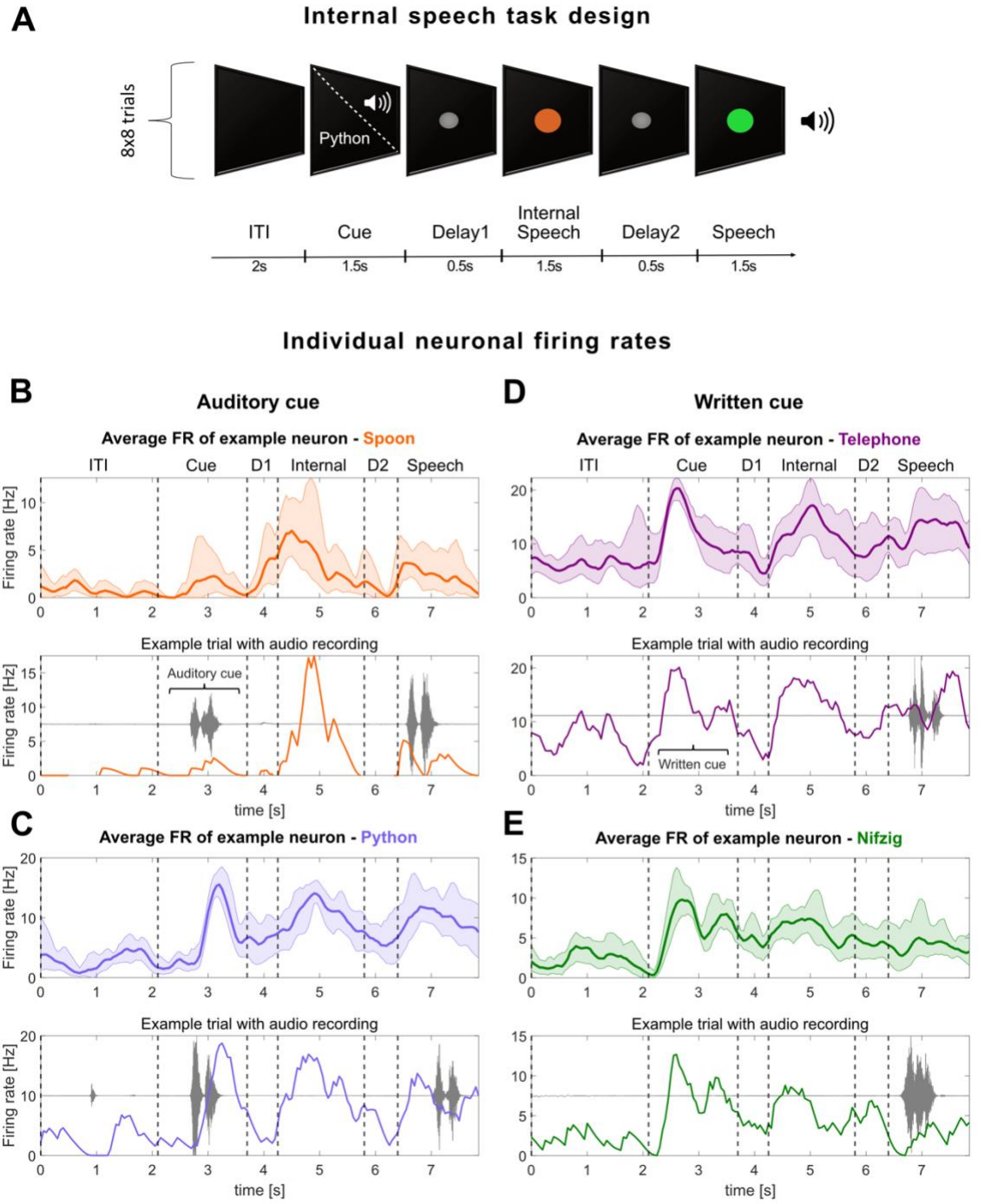

Figure 0-1 | Neurons in the supramarginal gyrus represent language processes. A) Written words and sounds were used to cue six words and two pseudowords in a tetraplegic participant. The "Audio cue" task was composed of an inter-trial interval (ITI), a cue phase during which the sound of one of the words was emitted from a speaker [between 842 – 1130ms], a first delay (D1), an internal speech phase, a second delay (D2) and a vocalized speech phase. The "Written cue" task was identical to the "Audio cue" task, apart that written words appeared on the screen for 1.5 seconds. Eight repetitions of eight words were performed per session day and per task. **B,C)** Example

smoothed firing rates of neurons tuned to four words in SMG during the "Audio cue", and **D,E)** the "Written cue" task. The top part of each word figure shows the average firing rate over eight trials (solid line: mean, shaded area: 95% bootstrapped confidence interval). The bottom part of each figure shows one out of eight example trials with associated audio amplitude (gray). Vertically dashed lines indicate the beginning of each phase.

These stereotypical activation patterns were evident at the single-trial level (**Figure 6-1A,B,C,D** bottom panel). When the auditory recording was overlaid with firing rates from a single trial, a heterogeneous neural response was observed (**Figure 6-S1A**), with some SMG neurons preceding or lagging peak auditory levels during vocalized speech. In contrast, neural activity from primary sensory cortex (S1) only modulated during vocalized speech, and produced stereotyped firing patterns regardless of the vocalized word (Figure 6-S1B).

Population activity represented selective tuning for individual words

Population analysis in SMG mirrored single-neuron patterns of activation, showing steep increases in tuning during the active task phases (Figure 6-2A). Tuning of a neuron to a word was determined by fitting a linear regression model to the firing rate in 50 ms time bins (methods). Representation of the auditory cue was lower compared to the written cue (**Figure 6-2B**, **Cue**). However, this difference was not observed for other task phases. The tuned population activity in S1 increased during vocalized speech, but not during the cue and internal speech phases (**Figure 6-S2A**).

To quantitatively compare activity between phases, we computed selectivity for individual words from the average FR in each task phase (**Figure 6-2B,C**). Tuning during the cue, internal speech, and vocalized speech phases was significantly higher compared to their preceding rest phases ITI, D1, and D2 (t-test, ** = $p < 0.01$, *** = $p < 0.001$). Representation for all words was observed in each phase, including pseudowords (Bindip and Nifzig) (**Figure 6-2C**).

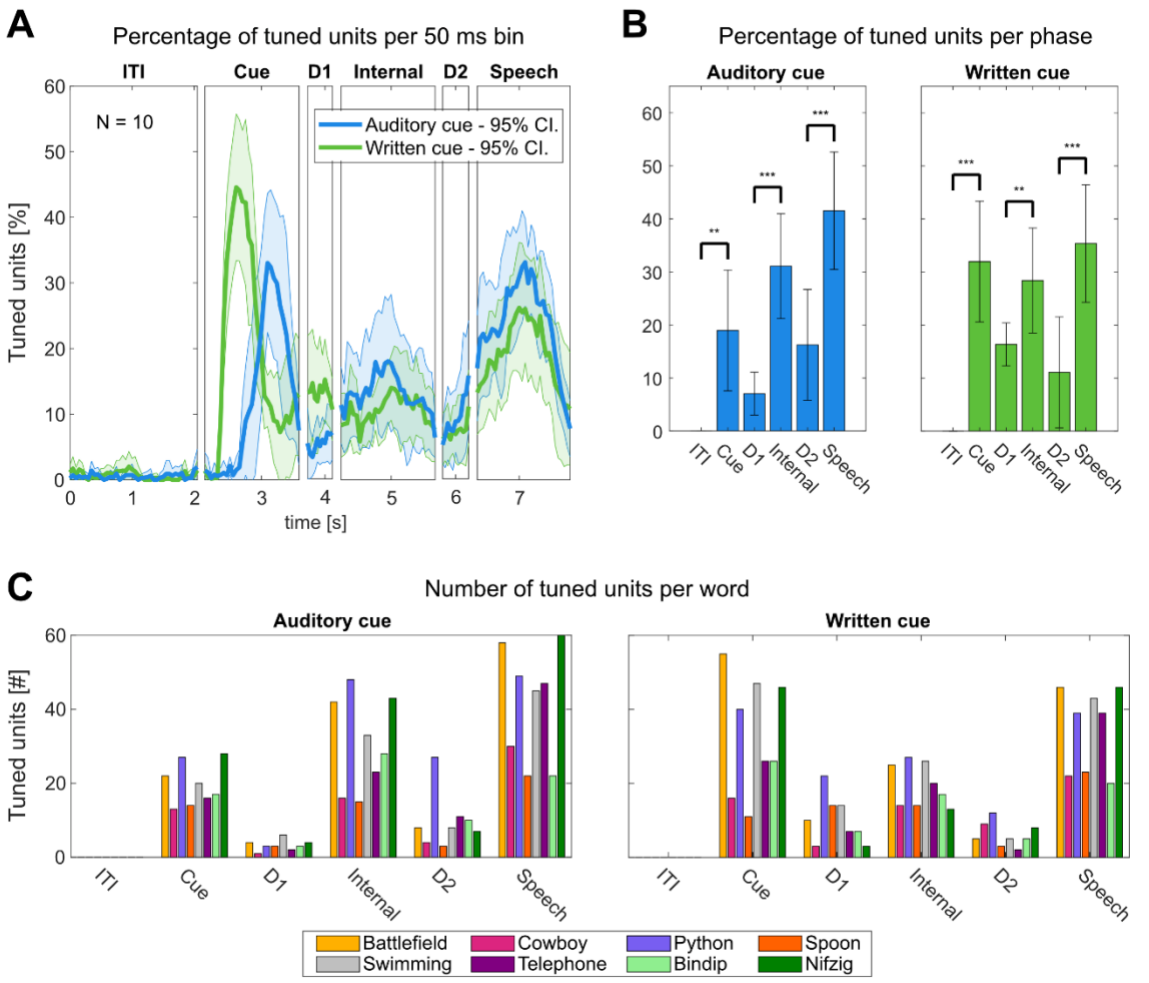

Figure 0-2 | Neuronal population activity modulates for individual words. A) Average percentage of tuned neurons to words in 50ms time bins in SMG over the trial duration for "Auditory cue" (blue) and "Written cue" (green) tasks (solid line: mean over 10 sessions, shaded area: 95% confidence interval). The written cue appeared on average 250ms earlier on the screen than the auditory sound. **B)** Average percentage of tuned neurons computed on firing rates per task phase, with 95% confidence interval over 10 sessions. Tuning during action phase (Cue, Internal, Speech) following rest phases (ITI, D1, D2) was significantly higher (t-test: **p < 0.01, ***p < 0.001). **C)** Number of neurons tuned to each individual word in each phase for the "Auditory cue" and "Written cue" tasks.

The neural population in SMG simultaneously represented several distinct aspects of language processing: temporal changes, input modality (auditory, written), and unique words from our vocabulary list (**Figure 6-3**). We used demixed Principal Components Analysis (dPCA) to decompose and analyze contributions of each individual component: Timing, Cue Modality, and Word.

The Timing component revealed temporal dynamics in SMG peaked during all active phases (**Figure 6-3A**)**.** In contrast, temporal S1 modulation peaked only during vocalized speech production, indicating a lack of synchronized lip and face movement of the participant during the other task phases (**Figure 6-3B**). While Cue Modality components were separable during the cue phase, they overlapped during subsequent phases. Thus, internal and vocalized speech representation may not be influenced by the cue modality. Pseudowords had similar separability to lexical words (**Figure 6-3C**). The explained variance between words was close to zero in S1.

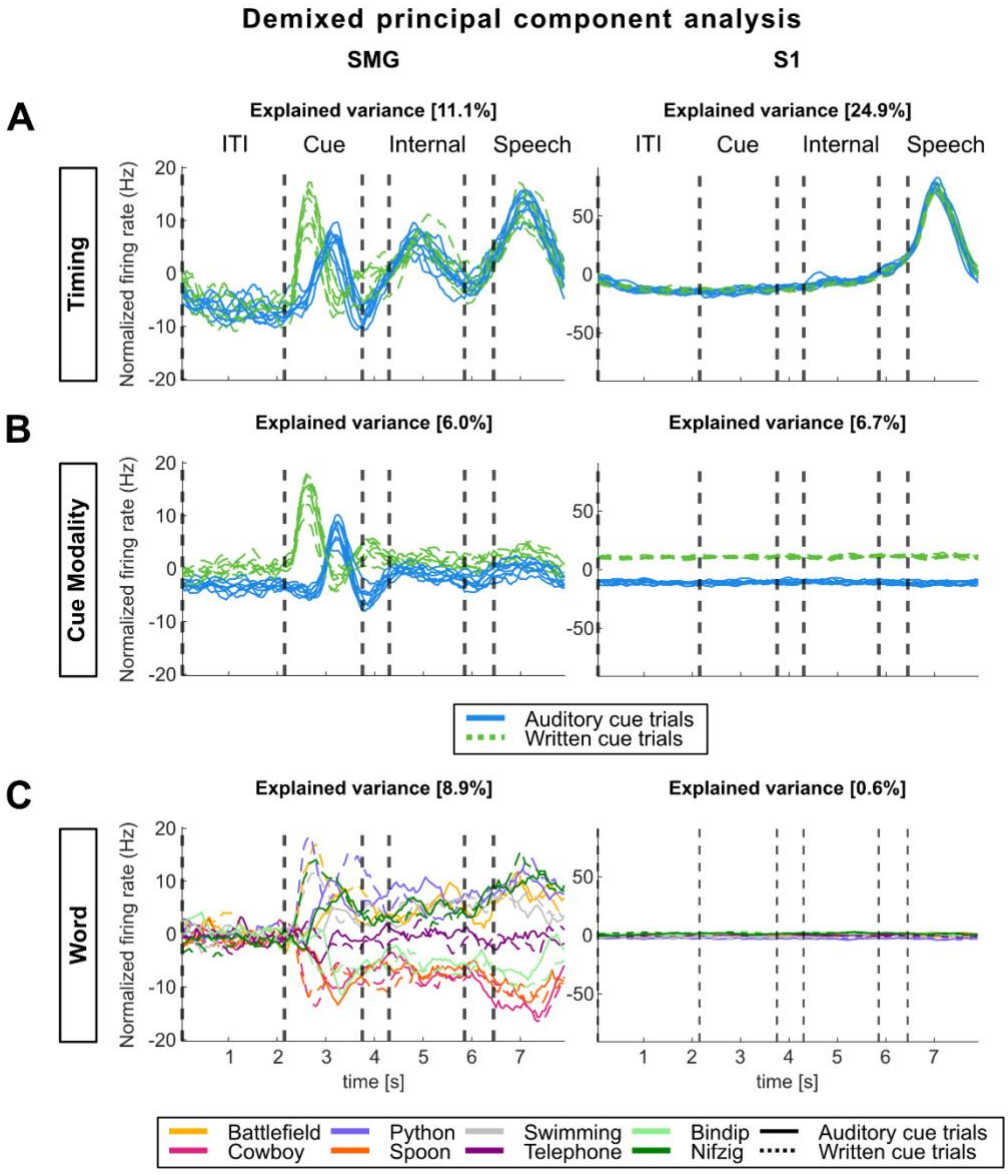

Figure 0-3 Demixed principal component analysis (dPCA) highlights SMG's involvement in language processing. dPCA was performed to investigate variance within three marginalizations: Timing, Cue Modality, and Word. Demixed principal components explaining the highest variance within each marginalization were plotted over time. A) The Timing marginalization demonstrates SMG modulation during cue, internal speech, and vocalized speech, while S1 only represents vocalized speech. B) The Cue Modality marginalization suggests internal and vocalized speech

representation in SMG are not affected by the cue modality. C) The Word marginalization shows high variability for different words in SMG, but near zero for S1.

Internal speech is highly decodable in SMG

Separable neural representations of both internal and vocalized speech processes implicate SMG as a rich source of neural activity for real-time speech BMI devices. All words in our vocabulary list were highly decodable, averaging 55% offline decoding and 84% online decoding from neurons during internal speech (Figure 6-4AB). Words spoken during the vocalized phase were also highly discriminable, averaging 74% offline (Figure 6-4A).

For offline analysis, trial data from both types of cues (auditory and written) were concatenated, since SMG activity was only differentiable between the type of cue during the cue phase (**Figure 6-2A**, **Figure 6-3B**). This resulted in 16 trials per condition. Features were selected via principal component analysis (PCA) on the training dataset, and principal components (PCs) which explained 95% of the variance were kept. A linear discriminant analysis (LDA) model was evaluated with leave-one-out cross-validation (CV). Significance was computed by comparing results to a null distribution (Methods).

Significant word decoding was observed during all phases, except during the ITI (**Figure 6-4A**, null distribution, turquoise - $p < 0.001$ = ***). Decoding accuracies were significantly higher in the cue, internal speech, and speech condition, compared to rest phases ITI, D1 and D2 (**Figure 6-4A**, t-test, black - $p < 0.001$ = ***). Significant cue phase decoding suggested that modality-independent linguistic representations were present early within the task (Leuthardt et al. 2012). Internal speech decoding averaged 55% offline, with the highest session at 72%, and a chance level of ~12.5% (**Figure 6-4A**, red line). Vocalized speech averaged even higher, at 74%. All words were highly decodable (**Figure 6-4C**). As suggested from our dPCA results, individual words were not significantly decodable from neural activity in S1 (**Figure 6-S2B**), indicating generalized activity for vocalized speech in the S1 arm region (**Figure 6-3C**).

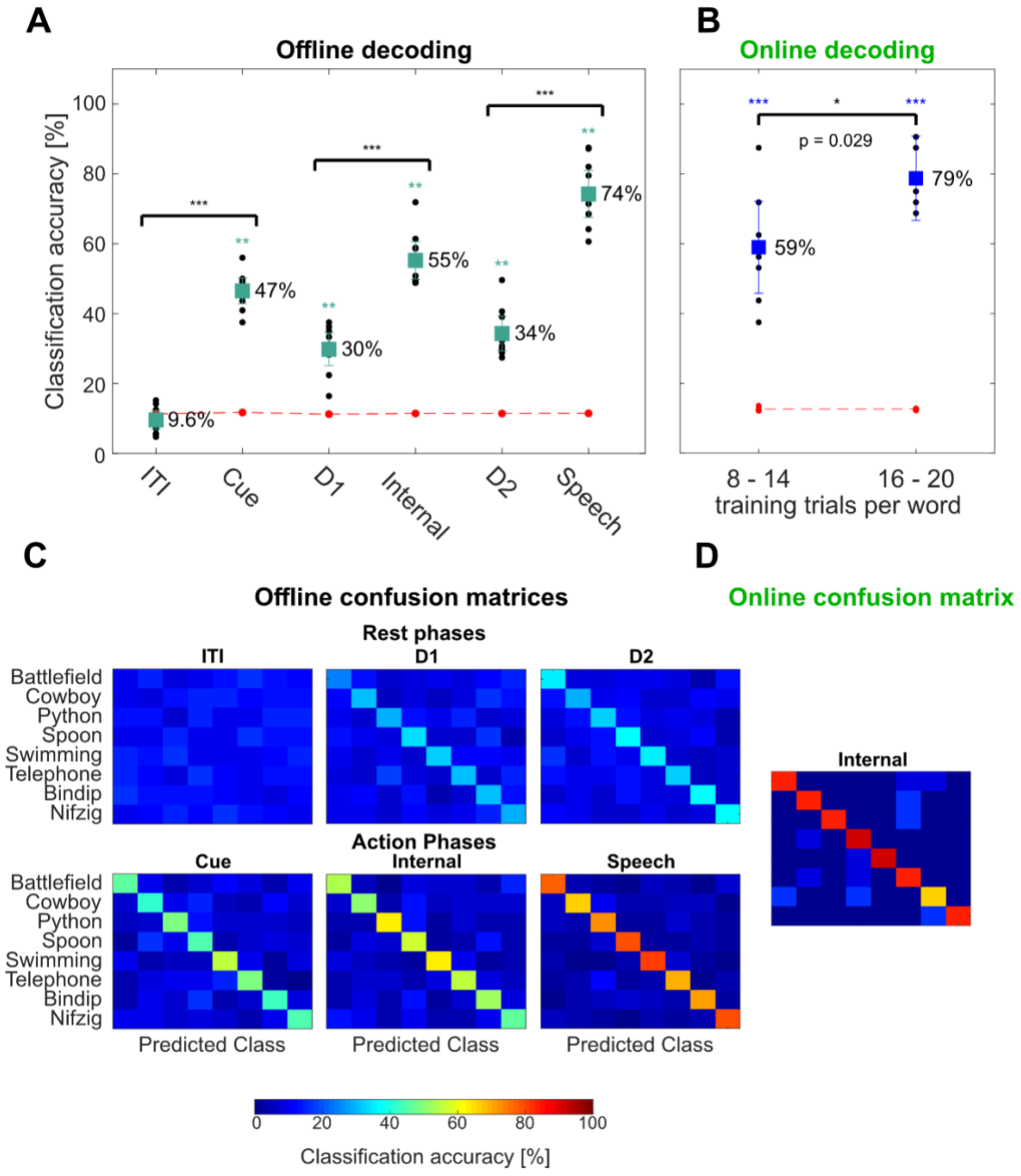

Figure 0-4 | Words can significantly be decoded during internal speech in SMG. A) Offline decoding accuracies: "Audio cue" and "Written cue" tasks data were combined for each individual session day, and leave one out cross-validation was performed (black dots). PCA was performed on the training data, a LDA model was constructed, and classification accuracies were plotted with 95% c.i., over the session means. Significance of classification accuracies was evaluated by comparing results to a shuffled distribution (averaged shuffle results = red dots, * = $p < 0.05$, ** = $p < 0.01$). Classification accuracies during action phases (Cue, Internal, Speech) following rest phases (ITI, D1, D2) were significantly higher (t-test: ***$p < 0.001$). **B)** Online decoding accuracies: Classification accuracies for internal speech were evaluated in a closed-loop internal speech BMI application. 72%, 88%, and 91% classification accuracies were achieved using respectively 8, 12, and 16 trials per word to train the classification model. **C)** Offline confusion matrix:

Confusion matrices for each of the different task phases were computed on the tested data, and averaged over all session days. **D)** Online confusion matrix: A confusion matrix was computed combing the three online runs.

High accuracy online speech decoder

We developed an online, closed-loop internal speech BMI with the highest achievable accuracy of 91% in an eight-word vocabulary (**Figure 6-4B**). A training dataset was generated using the written cue task, with 8 repetitions of each word. An LDA model was trained on the internal speech data of the training set, corresponding to only 96 seconds of neural data. The trained decoder predicted internal speech during the online task. During the online task, the vocalized speech phase was replaced with a feedback phase. The decoded word was shown in green if correctly decoded, and in red if wrongly decoded (**Video S1**). The classifier was retrained after each run of the online task, adding the newly recorded data. The first online run was trained on 8 repetitions per word and achieved 72% internal speech decoding. The second was trained on 12 trial repetitions and achieved 88% classification accuracy, while the third run was trained on 16 trial repetitions and achieved 91% classification accuracy. All words were well represented, illustrated by a confusion matrix (Figure 6-4D).

Shared representations between internal speech, written words, and vocalized speech

Several different language processes are engaged during the task: auditory comprehension or visual word recognition during the cue phase, and internal speech and vocalized speech production during the speech phases. It has been widely assumed each of these processes are part of a highly distributed network, involving multiple cortical areas (Indefrey and Levelt 2004). In this work, we observed significant representation of each of these processes in a common cortical region, SMG. To explore the relationships between each of these processes, we used cross-phase classification to identify the distinct and common neural codes. By training our classifier on the representation found in one phase (e.g. the cue phase) and testing the classifier on another phase (e.g. internal speech), we quantified generalizability of our models across neural activity of different language processes (**Figure 6-5**). The generalizability of a model to different task phases was evaluated through a t-test. No significant difference between classification accuracies indicated good generalization of the model, while significantly lower classification accuracies suggested poor generalization of the model.

The strongest shared neural representations were found between visual word recognition, internal speech, and vocalized speech (**Figure 6-5B**). A model trained on internal speech was highly generalizable to both vocalized speech and written cued words, evidence for a possible shared neural code (**Figure 6-5B**, Internal). In contrast, the model's performance was significantly lower when tested on data recorded in the auditory cue phase (**Figure 6-5A**, Internal, t-test, $p < 0.001$).

Additionally, a model trained on vocalized speech generalized equally well to internal speech and written words (**Figure 6-5B**, Speech). However, the model generalized significant better to internal speech than the representation found during the auditory cue (**Figure 6-5A**, Speech, Internal vs. Auditory Cue).

Neuronal representation of words at the single-neuron level was highly consistent between internal speech, vocalized speech, and written cue phases. A high percentage of neurons were not only active during the same task phases, but also were tuned to the same words (**Figure 6-5C,D**). 82-85% of neurons active during internal speech were also active during vocalized speech. In 53-56% of neurons, tuning was preserved between the internal speech and vocalized speech phases (**Figure 6-5C**). During the cue phase, 79% of neurons active during internal speech were also active during the written cue (**Figure 6-5D,** right). However, a significantly lower percentage of neurons (49%, t-test, $p < 0.01$ for all three other comparisons) were active during the auditory cue phase (**Figure 6-5D,** left). Similarly, 56% of neurons were tuned to the same words during the written cue phase as during the internal speech phase, while a significantly lower 21% of neurons were tuned to the same words during the auditory cue phase as during the internal speech phase.

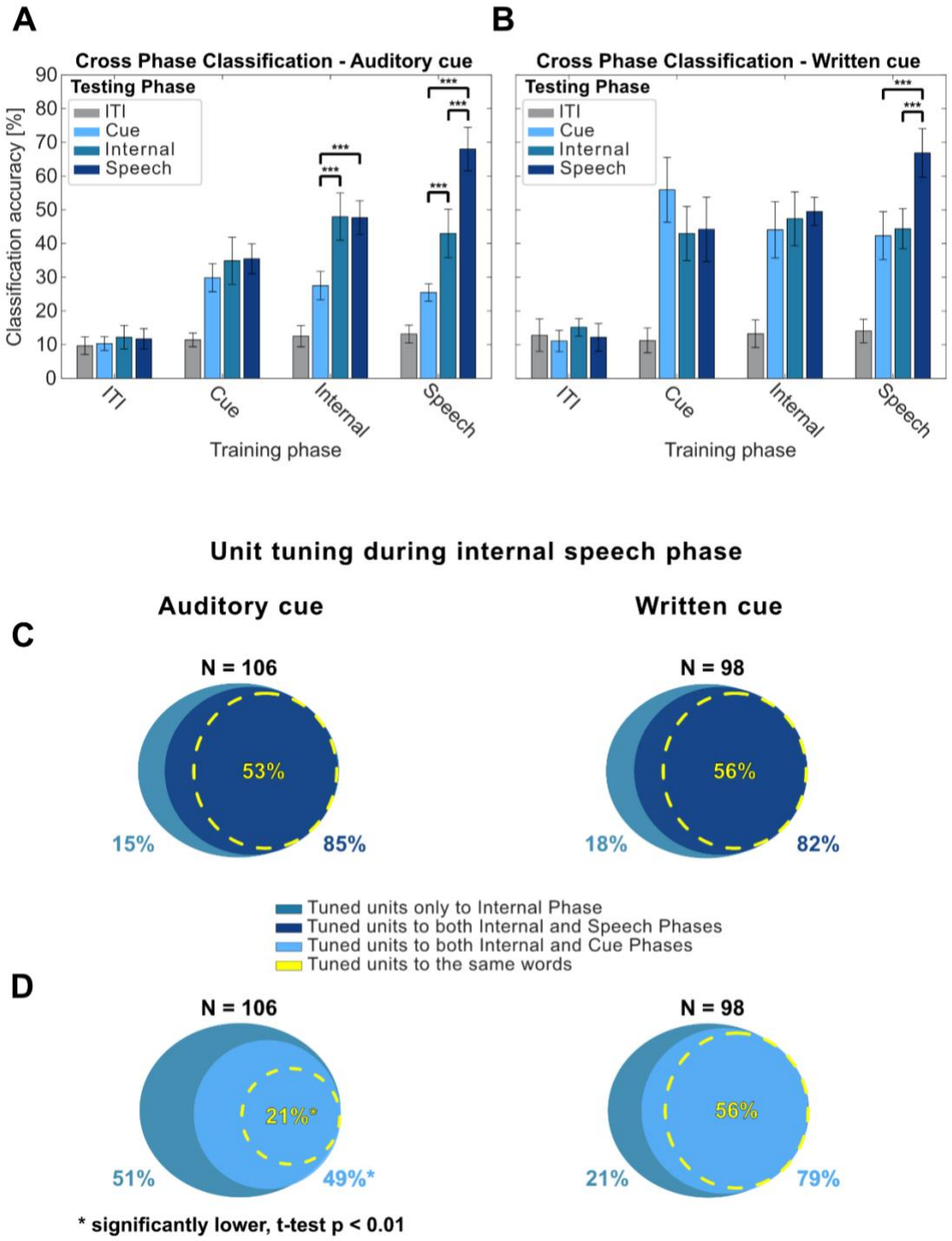

Figure 0-5 | Shared representations between internal speech, vocalized speech, and written word processing. A) Evaluating the overlap of shared information between different task phases in the "Auditory cue" task. For each of the 10 session days, cross-phase classification was performed. It consisted in training a model on a subset of data from one phase (e.g., Cue) and applying it on a subset of data from ITI, cue, internal, and speech phases. This analysis was performed separately for each task phase. PCA was performed on the training data, a LDA model was constructed, and classification accuracies were plotted with a 95% c.i. over session means. Significant differences in

performance between phases were evaluated between the 10 sessions (t-test, $p < 0.001$). For easier visibility, significant differences between ITI and other phases were not plotted. **B)** Same as A) for "Written cue" task. **C)** Percentage of neurons tuned during the internal speech phase which are also tuned during the vocalized speech phase. Neurons tuned during the internal speech phase were computed as in Figure 2B separately for each session day. From these, the percentage of neurons that was also tuned during vocalized speech was calculated. More than 80% of neurons during internal speech were also tuned during vocalized speech (85% in the "Auditory cue" task, 82% in the "Written cue" task). 53% of "Auditory cue" and 56% "Written cue" neurons also showed tuning to the same words during internal speech and vocalized speech phases. **D)** Percentage of neurons tuned during the internal speech phase which are also tuned during the cue phase. 79% of neurons tuned during internal speech were also tuned during the written cue phase (right side). A smaller 49% of neurons tuned during the internal speech phase were also tuned during the auditory cue phase. Percentages between Internal-Speech (85%, 82%) and Internal-Written cue (79%) were not significantly different. However, the overlapping Internal-Auditory cue percentage (49%) was significantly lower compared to the other overlaps (pairwise t-test, $p < 0.01$ for each comparison). 56% of neurons were tuned to the same words during the written cue phase and the internal speech phase. A significantly lower 21% of neurons were tuned to the same words during the auditory cue and the internal speech phase (pairwise t-test, $p < 0.01$ for each comparison).

Together with the cross-phase analysis, these results suggest strong shared neural representations between internal speech, vocalized speech, and written cue phase, both at the single-neuron and population level.

Robust decoding of multiple internal speech strategies within SMG

This shared neural representation between written, inner, and vocalized speech suggests that all three partly represent the same cognitive process or all cognitive processes share common neural features. While internal and vocalized speech have been shown to share common neural features (Pei, Leuthardt, et al. 2011b), similarities between internal speech and the written cue could have occurred through several different cognitive processes. For instance, the participant's observation of the written cue could have activated silent reading. This process has been self-reported as activating internal speech, which can involve "hearing" a voice, thus having an auditory component (Alderson-Day and Fernyhough 2015; Alderson-Day, Bernini, and Fernyhough 2017). However, the participant could also have mentally pictured an image of the written word while performing internal speech, involving visual imagination in addition to language processes. Both hypotheses could explain the high amount of shared neural representation between the written cue and the internal speech phases (**Figure 6-5B**).

We therefore compared two possible internal sensory strategies: a "sound imagination" strategy in which the participant imagined hearing the word, and a "visual imagination" strategy in which

the participant visualized the picture of the word (**Figure 6-S3A**). Both strategies were cued by each modality tested previously (auditory and written words).

Both employed strategies highly represented the four-word dataset (**Figure 6-S3B**, highest 94%, chance level: 25%). Furthermore, the participant described the "sound imagination" strategy as being easier and more similar to the internal speech condition of the first experiment. The participant's self-reported strategy suggests no visual imagination was performed during internal speech. Correspondingly, similarities between written cue and internal speech phases may stem from internal speech activation during the silent reading of the cue. These results suggest our speech BMI decoder is robust to multiple types of internal speech strategies.

Evidence of phonetic and semantic representation in SMG

Shared neural representations of the observed activation patterns could also have occurred if SMG encoded the semantic content of words. The native bilingual participant translated a four-word English vocabulary into Spanish, ensuring the semantic meaning of each word remained the same, while changing the phonetic content (**Figure 6-S4A**). Then, an English version of the task and a Spanish version of the task were run on three separate session days. Data were aggregated over languages and days, a support vector machine classifier model was used to evaluate the testing set with 8-fold cross-validation, and a confusion matrix was calculated. Both English and Spanish words were highly represented and separable. The offline model rarely confused words with the same semantic meaning, suggesting strong neural representation of phonetic content in SMG.

We then tested if homophones (words with the same pronunciation, but different meaning) would be decodable from SMG. Two groups of 3 homophones were tested: Scent, Sent, Cent, and Ware, Wear, Where. Data were analyzed as for the English/Spanish task. Interestingly, we found that words within the same homophone group were decodable above chance as well (**Figure 6-S5A**). Classification was higher during the cue phase than during internal and vocalized speech phases, suggesting neural data was more distinct when seeing the different spellings of the words than when pronouncing it. Nonetheless, decoding during internal and vocalized speech was well above chance (33.3%). Together, findings suggests both phonemes and semantics modulate neural activity in SMG.

Discussion

In this work, we demonstrated a robust decoder for internal and vocalized speech, capturing single-neuron activity from the supramarginal gyrus. A chronically implanted, speech-abled participant with tetraplegia was able to use an online, closed-loop internal speech BMI to achieve up to 91% classification accuracy with an eight-word vocabulary. Furthermore, high decoding required only 24 seconds of training data per word. Firing rates recorded from S1 showed generalized activation only during vocalized speech activity, but individual words were not classifiable. In SMG, shared neural representations between internal speech, the written cue, and vocalized speech suggest the occurrence of common processes. Robust control could be achieved using multiple internal speech strategies and for both English and Spanish vocabularies. Representation of pseudowords provided evidence for phonetic word encoding in SMG.

Single neurons in the supramarginal gyrus encode internal speech

We demonstrated internal speech decoding of six different words and two pseudowords in SMG. Single neurons increased their firing rates during internal speech (**Figure 6-1, S1**), which was also reflected at the population level (**Figure 6-2A,B**). Each word was well represented in the neuronal population (**Figure 6-2C, 3C**). Classification accuracy and tuning during the internal speech phase were significantly higher than during the previous delay phase (**Figure 6-2B, Figure 6-4A**). This evidence suggests we did not simply decode sustained activity from the cue phase, but activity generated by the participant performing internal speech. We obtained up to 72% offline classification accuracy of internal speech, and up to 91% during closed-loop online experiments (**Figure 6-4A**). These findings provide strong evidence for internal speech processing at the single-neuron level in SMG.

Neurons in primary somatosensory cortex are modulated by vocalized but not internal speech

Neural activity recorded from S1 served as a control for synchronized face and lip movements during internal speech. While vocalized speech robustly activated sensory neurons, no increase of baseline activity was observed during the internal speech phase or the auditory and written cue phases (**Figure 6-3, S1**). These results underline no synchronized movement inflated our decoding accuracy of internal speech (**Figure 6-S2**).

A previous imaging study achieved significant decoding of several different internal speech sentences performed by patients with mild ALS (Dash et al. 2020). Together with our findings, these results suggest a BMI speech decoder that does not rely on any movement may translate to communication opportunities for patients suffering from ALS and locked-in syndrome.

Different face activities are observable but not decodable in arm area of S1

The topographic representation of body parts in S1 has recently been found to be less rigid than previously thought. Generalized finger representation was found in a presumably S1 arm region of interest (ROI) (Rosenthal et al. 2022). Furthermore, a fMRI paper found observable face and lip activity in S1 leg and hand ROIs. However, differentiation between two lip actions was restricted to the face ROI (Muret et al. 2022). Correspondingly, we observed generalized face and lip activity in an predominantly S1 arm region (see Armenta Salas et al., 2018 for implant location) during vocalized speech (**Figure 6-3A**, **Figure 6-S1**, **Figure 6-S2A**). Recorded neural activity contained similar representations for different spoke words (**Figure 6-3C**) and was not significantly decodable (**Figure 6-S2B).**

Shared neural representations between internal and vocalized speech

The extent to which internal and vocalized speech generalize is still debated (Cooney, Folli, and Coyle 2018; Perrone-Bertolotti et al. 2014; Alderson-Day and Fernyhough 2015), and depends on the investigated brain area (Pei, Leuthardt, et al. 2011a; Soroush et al. 2022). In this work, we found on average stronger representation for vocalized (74%) than internal speech (**Figure 6-4A**, 55%). Additionally, cross-phase decoding of vocalized speech from models trained on data during internal speech resulted in comparable classification accuracies to those of internal speech (**Figure 6-5A,B,** Internal). However, in certain contexts we observed identical or better classification accuracy during internal speech than vocalized speech (**Figure 6-4A,** Online decoding). Better decoding of individual internal words than vocalized words was also observed during the phonetic task (**Figure 6-S4**, Swimming). Most neurons tuned during internal speech were also tuned to the same words during vocalized speech (53-56%, **Figure 6-5C**). However, some neurons were only tuned during internal speech, or to different words. These observations also applied to firing rates of individual neurons. Here, we observed neurons that had higher peak rates during the internal speech phase than the vocalized speech phase (**Figure 6-1**: Spoon, Nifzig;

Figure 6-S1: Swimming, Cowboy). Together, these results suggest SMG does not present a strict nested hierarchy of internal and imagined speech, where all channels involved during imagined speech are a subset of channels involved during vocalized speech (Soroush et al. 2022).

Similar observations were made when comparing internal speech processes to visual word processes. 79% of neurons were active both in the internal speech phase and the written cue phase, and 56% preserved the same tuning (**Figure 6-5D**, Written Cue). Additionally, high cross-decoding between both phases was observed (**Figure 6-5B**).

Shared representation between speech and written cue presentation

Observation of a written cue may engage a variety of cognitive processes, such as visual feature recognition, semantic understanding and/or related language processes, many of which modulate similar cortical regions as speech (Leuthardt et al. 2012). Studies have found silent reading can evoke internal speech; it can be modulated by presumed author's speaking speed, voice familiarity, or regional accents (Perrone-Bertolotti et al. 2014; Alderson-Day and Fernyhough 2015; Alderson-Day, Bernini, and Fernyhough 2017; Alexander and Nygaard 2008; Filik and Barber 2011; Lœvenbruck et al. 2005). During silent reading of a cued sentence with a neutral vs. increased prosody (madeleine brought me vs. MADELEINE brought me), one study in particular found increased left SMG activation correlated with the intensity of the produced inner speech (Lœvenbruck et al. 2005).

Our data demonstrated high cross-phase decoding accuracies between both written cue and speech phases (**Figure 6-5B**). Due to substantial shared neural representation, we hypothesize the participant's silent reading during the presentation of the written cue may have engaged internal speech processes. However, this same shared representation could have occurred if visual processes were activated in the internal speech phase. For instance, the participant could have performed mental visualization of the written word instead of generating an internal monologue, as the subjective perception of internal speech may vary between individuals.

Investigating internal speech strategies

In a separate experiment, the participant was prompted to execute different mental strategies during the internal speech phase, consisting of "sound imagination" or "visual word imagination"

(**Figure 6-S3A**). We found robust decoding during the internal strategy phase, regardless of which mental strategy was performed (**Figure 6-S3B**). The participant reported the sound strategy was easier to execute than the visual strategy. Furthermore, the participant reported the sound strategy was more similar to the internal speech strategy employed in prior experiments. This self-report suggests the patient did not perform visual imagination during the internal speech task. Therefore, shared neural representation between internal and written word phases during the internal speech task may stem from silent reading of the written cue. Since multiple internal mental strategies are decodable from SMG, future patients could have flexibility with their preferred strategy. For instance, people with a strong visual imagination may prefer performing visual word imagination.

Phonetic and semantic processing in SMG

Since substantial shared neuronal representation was observed across the different task phases, semantic information in SMG might be responsible. Our native bilingual participant performed the same internal speech experiment using a four-word English and four-word Spanish vocabulary of words with the same meaning. If semantics were encoded, we hypothesized a classification model trained on English and Spanish words would confuse words of the same meaning. Instead, we observed stronger representation of phonetic information, including robust representation of pseudowords with no semantic content (**Figure 6-2C**, **Figure 6-3C, Figure 6-S4**). Our findings in SMG agree with previous literature reports of stronger SMG representation for phonetic rather than semantic decisions on words (Oberhuber et al. 2016; C. J. Price et al. 1997; Seghier et al. 2004; Sliwinska et al. 2012b). To test if word meaning also modulate SMG activity, two groups of homophones were tested. Interestingly, these words were distinguishable from each other as well (**Figure 6-S5**), suggesting both phonetic and semantics modulate SMG activity.

Audio contamination in decoding result

Prior studies examining neural representation of attempted or vocalized speech must potentially mitigate acoustic contamination of electrophysiological brain signals during speech production (Roussel et al. 2020). During internal speech production, no detectable audio was captured by the audio equipment or noticed by the researchers in the room. In the rare cases the participant spoke during internal speech (3 trials), the trial was removed. Furthermore, if audio had contaminated

the neural data during the auditory cue or vocalized speech, we would have likely observed significant decoding in all channels. However, no significant classification was detected in S1 channels. We therefore conclude acoustic contamination did not artificially inflate observed classification accuracies during vocalized speech in SMG.

Impact on BMI applications

In this work, an online internal speech BMI achieved high-performance from single-neuron activity in SMG. The online decoders were trained on as little as eight repetitions of 1.5 seconds per word, demonstrating meaningful classification accuracies can be obtained with merely a few minutes' worth of training data per day. This proof-of-concept suggests SMG may be able to represent a much larger internal vocabulary. By building models on internal speech directly, our results may translate to people who cannot vocalize speech or are completely locked in. Recently, (Metzger et al. 2022) demonstrated a BMI speller that worked by decoding attempted speech of the letters of the NATO alphabet, and using those to construct sentences. Scaling our vocabulary to that size could allow for unrestricted internal speech communication.

To summarize, we demonstrate SMG as a promising candidate to build an internal brain-machine speech device. Different internal speech strategies were decodable from SMG, allowing patients to use the methods and languages with which they are most comfortable. We found evidence for phonetic representation during internal and vocalized speech. Adding to previous findings indicating grasp decoding in SMG (Wandelt et al. 2022a), we pose SMG as a multipurpose brain-machine interface area.

Acknowledgements

We wish to thank L. Bashford, and I. Rosenthal for helpful discussions and data collection. We wish to thank our study participant FG for his dedication to the study which made this work possible. This research was supported by the NIH National Institute of Neurological Disorders and Stroke Grant U01: U01NS098975 and U01: U01NS123127 (S.K.W, D.B., K.P., C.L., and R.A.A.), and by the T&C Chen Brain-machine Interface center (S.K.W., D.B., R.A.A.).

Author contributions

S.K.W., D.B., and R.A.A. designed the study. S.K.W. developed the experimental tasks and analyzed the results. S.K.W., D.B., and R.A.A. interpreted the results and wrote the paper. K.P. coordinated regulatory requirements of clinical trials. C.L. and B.L. performed the surgery to implant the recording arrays.

Supplemental Figures

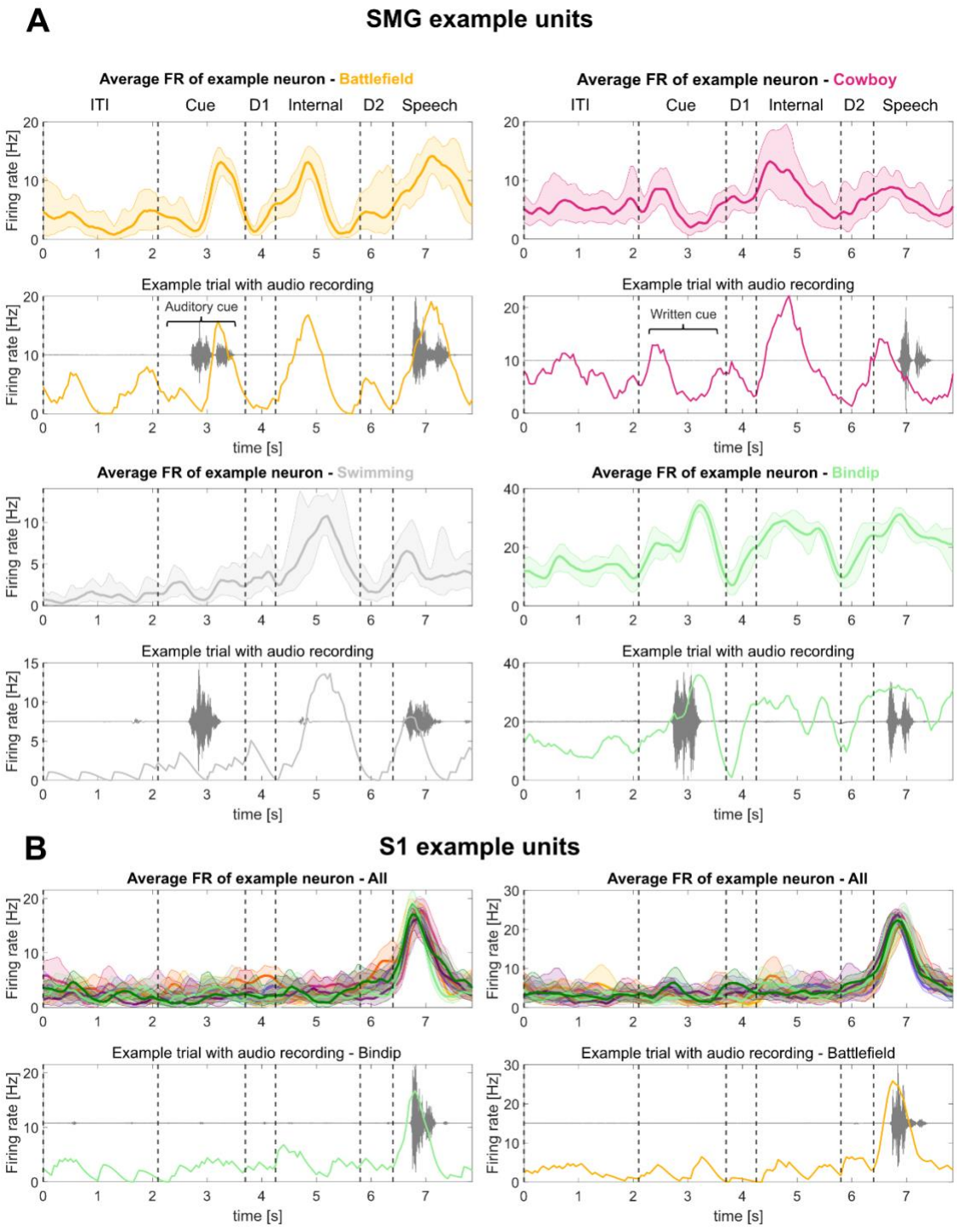

Figure S6-1 | SMG shows firing rate modulation during cue, internal speech, and vocalized speech; S1 shows firing rate modulation during vocalized speech. A) Additional example smoothed firing rates of neurons tuned to four words in SMG during the "Auditory cue" and the "Written cue" task. The top part of each word figure shows the average firing rate over eight trials (solid line: mean, shaded area: 95% bootstrapped confidence interval). The bottom part of each figure shows an example trial with associated audio amplitude (gray). Vertically dashed lines indicate the beginning of each phase. **B)** Example smoothed firing rates for S1 over task duration. Tuning of a neuron to all words simultaneously was shown to emphasize generalized speech activity to vocalized words.

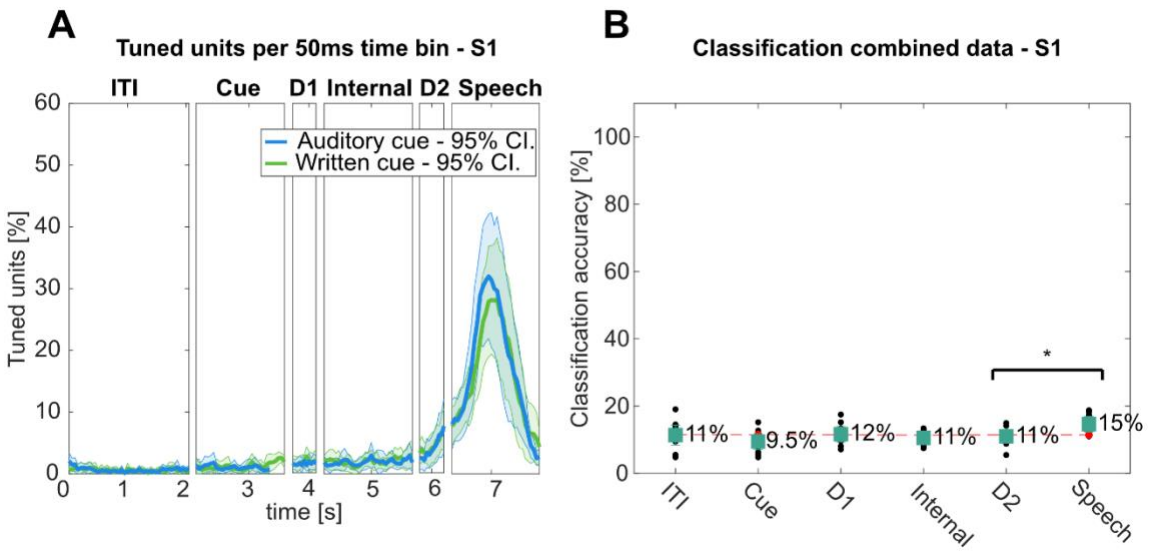

Figure S6-2 | S1 shows generalized word activity during vocalized speech. A) Average percentage of tuned neurons to words in 50ms time bins in S1 over the trial duration for "Auditory cue" (blue) and "Written cue" (green) tasks (solid line: mean over 10 sessions, shaded area: 95% confidence interval). **B)** "Auditory cue" and "Written cue" tasks data were combined for each individual session day, and leave-one-out cross-validation was performed (black dots). PCA was performed on the training data, a LDA model was constructed, and results were plotted with 95% c.i., of the session means. Significance of classification accuracies was evaluated by comparing results to a shuffled distribution (averaged shuffle results = red dots). No classification accuracy was significant. However, classification accuracy during vocalized speech was significantly higher than during the previous delay period (t-test: *$p < 0.05$). These results show while lip and face activity are represented in putative arm area in S1, it is not significantly decodable.

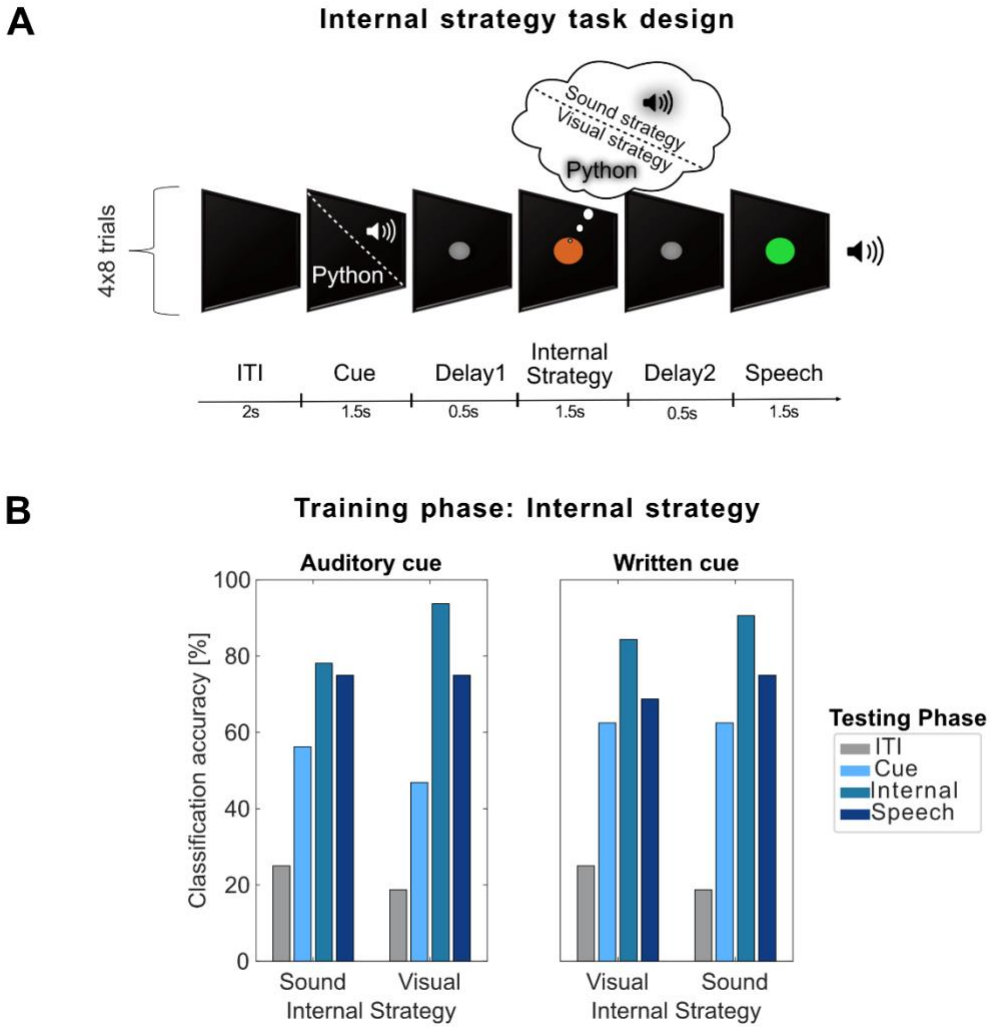

Figure S6-3 | Different internal speech strategies are represented in SMG. A) The task was designed to vary the internal strategy the participant was performing during the internal speech phase. Two internal speech strategies were tested: a sound imagination and a visual imagination strategy. For the "sound imagination" strategy, the participant was instructed to imagine the sound of the word. For the "visual imagination" strategy, the participant was instructed to perform mental visualization of the written word. To test if the cue modality (auditory or written) could influence the internal strategy, each internal strategy was run once with an auditory cue, and once with a written cue, resulting in four different task versions (Auditory/Sound, Auditory/Visual, Written/Sound, Written/Visual; see Methods). A subset of four words was used for this experiment. **B)** Cross-phase classification was performed by training the model on a subset of data from one phase (e.g. Cue) and applying it on a subset of data from each phase. This analysis was performed separately for each phase, and for each of the four task versions. Plotted here are the results when training on the internal speech phase, and evaluating it on ITI, Cue, Internal, and Speech phases. High classification accuracies (up to 94%) while performing the internal strategy were achieved using both visual and sound imagination strategy.

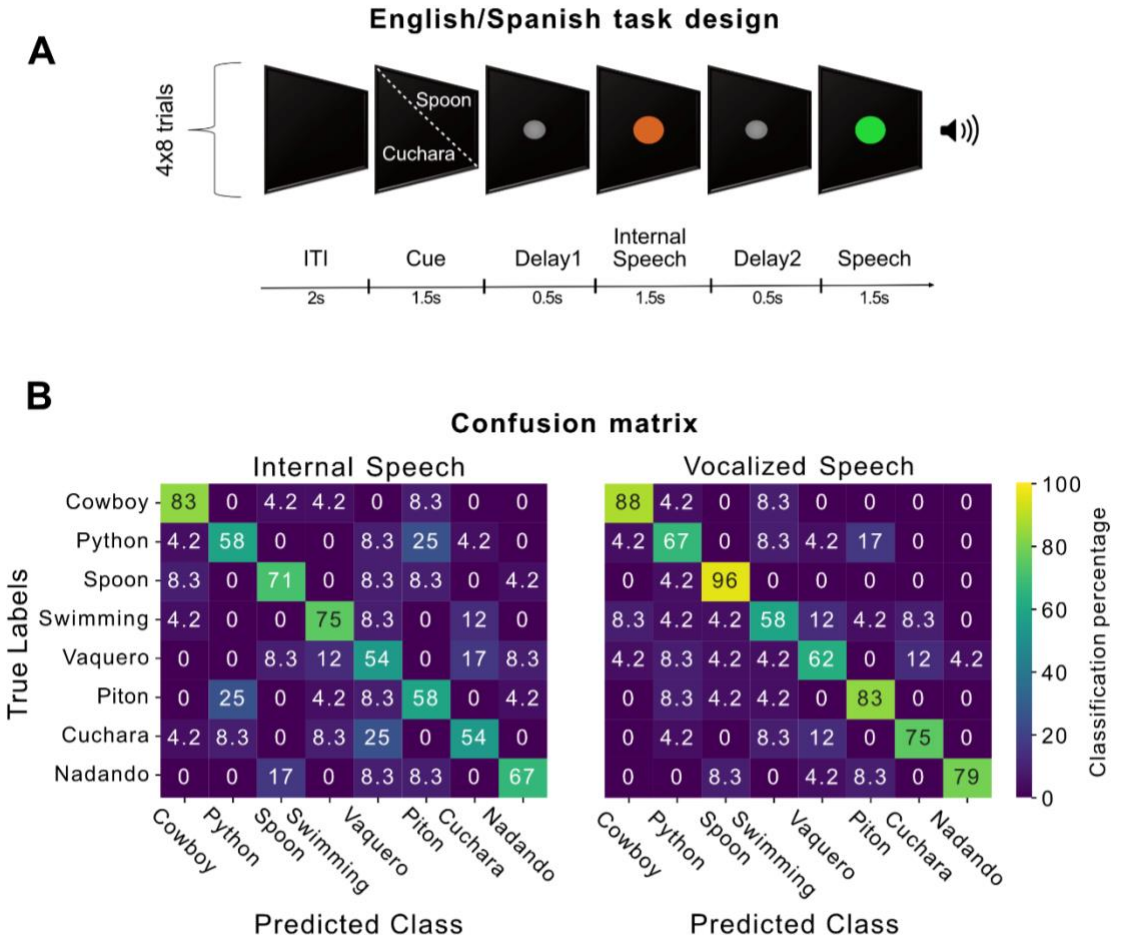

Figure S6-4 | SMG encodes sound phonetics. A) The task was designed to probe if SMG encodes stronger the semantics (e.g. the concept of a word) or the phonetics of a word. The bilingual Spanish participant translated four English words into their Spanish counterparts, to ensure the semantic meaning of the word remained the same. The Spanish version of the task was composed of the translated words. On three session days, an English version and a Spanish version of the task were run, resulting in 24 trials per word. The task design remained identical to task 1. **B)** Data was concatenated over session days and languages. A support-vector class model was trained and evaluated with 8-fold cross validation on multiunit channel activity. The confusion matrix during the internal speech phase (left) and vocalized speech phase (right) was calculated. Results indicate the classifier rarely got confused between words with the same meaning, suggesting SMG processes phonetics stronger than semantics both during internal and vocalized speech.

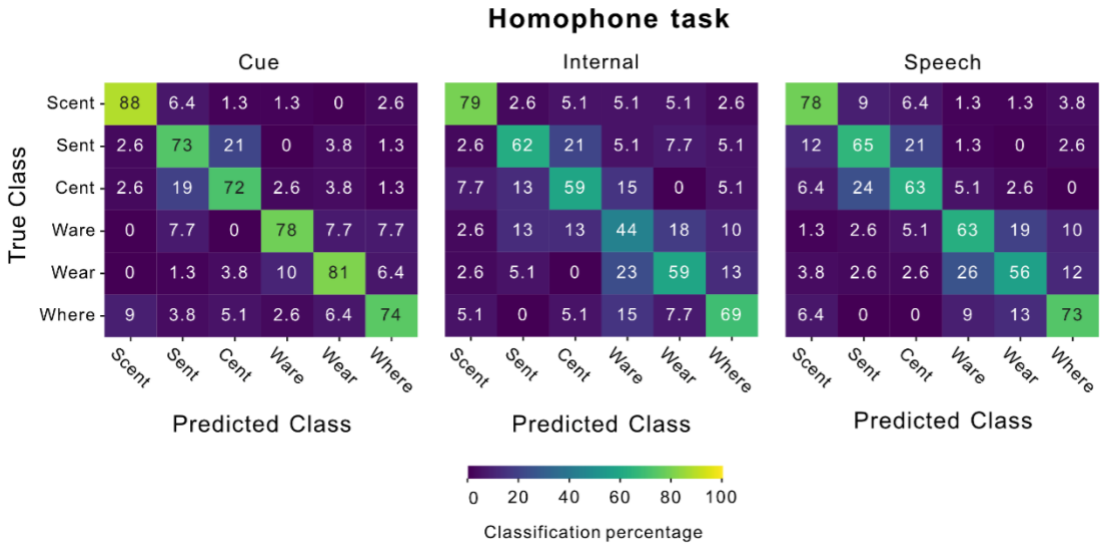

Figure S6-5 | SMG encodes word meaning. The task was designed to probe if SMG represents the semantic meaning of words. On four session days, two groups of three homophones were run, resulting in 78-80 trials per word. The task design remained identical to task 1. Data was concatenated over session days. A support-vector class model was trained and evaluated with 8-fold cross validation on multiunit channel activity. The confusion matrix during the cue phase, internal speech phase, and vocalized speech phase was calculated. Results indicate the classifier was able to decode words with same phonetics above chance, suggesting neural activity in SMG is also modulated by word meaning.

Chapter 7:
Conclusion

The ability to move freely and to connect with others through communication is invaluable for human independence. For individuals who have lost these crucial skills, regaining them is their main priority (Anderson 2004; Hecht et al. 2002). In this thesis, we demonstrated novel applications of brain-machine interfaces for patients suffering from paralysis and speech impairments. Our work builds on previous research indicating that premotor and posterior parietal areas are involved in both reach and grasp movements and language processes (Andersen et al. 2014a; Andersen, Aflalo, and Kellis 2019; Tyson Aflalo et al. 2015; Klaes et al. 2015; T. Aflalo et al. 2020; Oberhuber et al. 2016).

Using 4 mm x 4 mm multielectrode Utah arrays, we recorded populations of single neurons and found that the supramarginal gyrus (SMG), the ventral premotor cortex (PMv), and primary somatosensory cortex (S1) significantly encode motor imagery of grasping. These findings strengthen the idea of BMI target sites in higher level brain areas, that are involved in both grasp planning and execution. We further investigated the cognitive processes underlying neural activity during the cue phase of grasping and found a transition from cue modality-dependent to cue modality-independent grasp representation in SMG, the anterior intraparietal cortex (AIP) and PMv. Our findings suggest SMG integrates audio, written, and image cues modalities, but more similarly represents audio and written cues, indicating language involvement. We confirmed this hypothesis by demonstrating that SMG encodes both grasp and speech processes and engages different motor plans for each. For BMI applications, these results suggest that a classifier could be trained to represent both processes simultaneously, thus reducing the number of required implantable target sites. Lastly, we showed SMG involvement during internal speech and identified it as a promising candidate for developing an internal brain-machine speech device, that decoded eight words with high accuracy. Different internal speech strategies as well as English and Spanish words modulated SMG activity, suggesting participants can use the methods and languages with which they are most comfortable. Our findings also provide evidence for phonetic as well as semantic representation during internal and vocalized speech. These results are proof-of-concept that

internal speech BMIs can be constructed from brain signals obtained from the PPC. To increase the effectiveness of the current device, increasing the vocabulary size, the amount of training data, and improving the decoding algorithm are all important considerations. Constructing a speller by decoding words of the NATO alphabet (Moses et al. 2021), or phonemes paired with a language model (F. Willett et al. 2023) could allow to scale the vocabulary substantially.

In this work, we provided proof-of-concept for grasp and speech BMI applications from the posterior parietal cortex. An important future step to validate their effectiveness and potential benefits is by replicating the results in additional participants. This task can be challenging, as invasiveness of the procedure, specific medical conditions, and limited access to resources can prevent potential participants from enrolling in studies of this caliber. Additionally, each participant tends to have unique brain structures and individual variabilities. Nonetheless, the findings in this thesis are promising for developing more natural and intuitive BMI applications.

Overall, this thesis highlights the potential of BMI applications to improve the quality of life for those affected by the loss of essential skills for human independence.

BIBLIOGRAPHY

Aflalo, T., C. Y. Zhang, E. R. Rosario, N. Pouratian, G. A. Orban, and R. A. Andersen. 2020. "A Shared Neural Substrate for Action Verbs and Observed Actions in Human Posterior Parietal Cortex." *Science Advances* 6 (43): eabb3984. https://doi.org/10.1126/sciadv.abb3984.

Aflalo, Tyson, Spencer Kellis, Christian Klaes, Brian Lee, Ying Shi, Kelsie Pejsa, Kathleen Shanfield, et al. 2015. "Decoding Motor Imagery from the Posterior Parietal Cortex of a Tetraplegic Human." *Science (New York, N.Y.)* 348 (6237): 906–10. https://doi.org/10.1126/science.aaa5417.

Alderson-Day, Ben, Marco Bernini, and Charles Fernyhough. 2017. "Uncharted Features and Dynamics of Reading: Voices, Characters, and Crossing of Experiences." *Consciousness and Cognition* 49 (March): 98–109. https://doi.org/10.1016/j.concog.2017.01.003.

Alderson-Day, Ben, and Charles Fernyhough. 2015. "Inner Speech: Development, Cognitive Functions, Phenomenology, and Neurobiology." *Psychological Bulletin* 141 (5): 931–65. https://doi.org/10.1037/bul0000021.

Alexander, Jessica D., and Lynne C. Nygaard. 2008. "Reading Voices and Hearing Text: Talker-Specific Auditory Imagery in Reading." *Journal of Experimental Psychology: Human Perception and Performance* 34 (2): 446–59. https://doi.org/10.1037/0096-1523.34.2.446.

Andersen, Richard A. 2019. "Machines That Translate Wants into Actions." Scientific American. 2019. https://doi.org/10.1038/scientificamerican0419-24.

Andersen, Richard A., Tyson Aflalo, and Spencer Kellis. 2019. "From Thought to Action: The Brain–Machine Interface in Posterior Parietal Cortex." *Proceedings of the National Academy of Sciences* 116 (52): 26274–79. https://doi.org/10.1073/pnas.1902276116.

Andersen, Richard A., and Christopher A. Buneo. 2002. "Intentional Maps in Posterior Parietal Cortex." *Annual Review of Neuroscience* 25 (1): 189–220. https://doi.org/10.1146/annurev.neuro.25.112701.142922.

Andersen, Richard A., Spencer Kellis, Christian Klaes, and Tyson Aflalo. 2014a. "Toward More Versatile and Intuitive Cortical Brain Machine Interfaces." *Current Biology : CB* 24 (18): R885–97. https://doi.org/10.1016/j.cub.2014.07.068.

———. 2014b. "Toward More Versatile and Intuitive Cortical Brain Machine Interfaces." *Current Biology : CB* 24 (18): R885–97. https://doi.org/10.1016/j.cub.2014.07.068.

Anderson, Kim D. 2004. "Targeting Recovery: Priorities of the Spinal Cord-Injured Population." *Journal of Neurotrauma* 21 (10): 1371–83. https://doi.org/10.1089/neu.2004.21.1371.

Andric, Michael, Ana Solodkin, Giovanni Buccino, Susan Goldin-Meadow, Giacomo Rizzolatti, and Steven L. Small. 2013. "Brain Function Overlaps When People Observe Emblems, Speech, and Grasping." *Neuropsychologia* 51 (8): 1619–29. https://doi.org/10.1016/j.neuropsychologia.2013.03.022.

Angrick, Miguel, Christian Herff, Emily Mugler, Matthew C. Tate, Marc W. Slutzky, Dean J. Krusienski, and Tanja Schultz. 2018. "Speech Synthesis from ECoG Using Densely Connected 3D Convolutional Neural Networks." Preprint. Neuroscience. https://doi.org/10.1101/478644.

Angrick, Miguel, Maarten C. Ottenhoff, Lorenz Diener, Darius Ivucic, Gabriel Ivucic, Sophocles Goulis, Jeremy Saal, et al. 2021. "Real-Time Synthesis of Imagined Speech Processes from Minimally

Invasive Recordings of Neural Activity." *Communications Biology* 4 (1): 1–10. https://doi.org/10.1038/s42003-021-02578-0.

Anumanchipalli, Gopala K., Josh Chartier, and Edward F. Chang. 2019. "Speech Synthesis from Neural Decoding of Spoken Sentences." *Nature* 568 (7753): 493–98. https://doi.org/10.1038/s41586-019-1119-1.

Armenta Salas, Michelle, Luke Bashford, Spencer Kellis, Matiar Jafari, HyeongChan Jo, Daniel Kramer, Kathleen Shanfield, et al. 2018. "Proprioceptive and Cutaneous Sensations in Humans Elicited by Intracortical Microstimulation." Edited by Ranulfo Romo. *ELife* 7 (April): e32904. https://doi.org/10.7554/eLife.32904.

Bashford, Luke, Isabelle Rosenthal, Spencer Kellis, Kelsie Pejsa, Daniel Kramer, Brian Lee, Charles Liu, and Richard A. Andersen. 2021. "The Neurophysiological Representation of Imagined Somatosensory Percepts in Human Cortex." *The Journal of Neuroscience* 41 (10): 2177–85. https://doi.org/10.1523/JNEUROSCI.2460-20.2021.

Baumann, Markus A., Marie-Christine Fluet, and Hansjörg Scherberger. 2009. "Context-Specific Grasp Movement Representation in the Macaque Anterior Intraparietal Area." *Journal of Neuroscience* 29 (20): 6436–48. https://doi.org/10.1523/JNEUROSCI.5479-08.2009.

Binder, Jeffrey R. 2017. "Current Controversies on Wernicke's Area and Its Role in Language." *Current Neurology and Neuroscience Reports* 17 (8): 58. https://doi.org/10.1007/s11910-017-0764-8.

Bocquelet, Florent, Thomas Hueber, Laurent Girin, Christophe Savariaux, and Blaise Yvert. 2016. "Real-Time Control of an Articulatory-Based Speech Synthesizer for Brain Computer Interfaces." *PLOS Computational Biology* 12 (11): e1005119. https://doi.org/10.1371/journal.pcbi.1005119.

Bonini, Luca, Stefano Rozzi, Francesca Ugolotti Serventi, Luciano Simone, Pier F. Ferrari, and Leonardo Fogassi. 2010. "Ventral Premotor and Inferior Parietal Cortices Make Distinct Contribution to Action Organization and Intention Understanding." *Cerebral Cortex* 20 (6): 1372–85. https://doi.org/10.1093/cercor/bhp200.

Branco, Mariana P., Zachary V. Freudenburg, Erik J. Aarnoutse, Martin G. Bleichner, Mariska J. Vansteensel, and Nick F. Ramsey. 2017. "Decoding Hand Gestures from Primary Somatosensory Cortex Using High-Density ECoG." *NeuroImage* 147 (February): 130–42. https://doi.org/10.1016/j.neuroimage.2016.12.004.

Brandman, David M., Sydney S. Cash, and Leigh R. Hochberg. 2017. "Review: Human Intracortical Recording and Neural Decoding for Brain-Computer Interfaces." *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society* 25 (10): 1687–96. https://doi.org/10.1109/TNSRE.2017.2677443.

Bruni, Stefania, Valentina Giorgetti, Leonardo Fogassi, and Luca Bonini. 2017. "Multimodal Encoding of Goal-Directed Actions in Monkey Ventral Premotor Grasping Neurons." *Cerebral Cortex* 27 (1): 522–33. https://doi.org/10.1093/cercor/bhv246.

Buchwald, Mikolaj, Łukasz Przybylski, and Gregory Króliczak. 2018. "Decoding Brain States for Planning Functional Grasps of Tools: A Functional Magnetic Resonance Imaging Multivoxel Pattern Analysis Study." *Journal of the International Neuropsychological Society* 24 (10): 1013–25. https://doi.org/10.1017/S1355617718000590.

Carpaneto, J., M. A. Umiltà, L. Fogassi, A. Murata, V. Gallese, S. Micera, and V. Raos. 2011. "Decoding the Activity of Grasping Neurons Recorded from the Ventral Premotor Area F5 of the

Macaque Monkey." *Neuroscience* 188 (August): 80–94. https://doi.org/10.1016/j.neuroscience.2011.04.062.

Chaudhary, U., N. Birbaumer, and M. R. Curado. 2015. "Brain-Machine Interface (BMI) in Paralysis." *Annals of Physical and Rehabilitation Medicine* 58 (1): 9–13. https://doi.org/10.1016/j.rehab.2014.11.002.

Choi, Jong-ryul, Seong-Min Kim, Rae-Hyung Ryu, Sung-Phil Kim, and Jeong-woo Sohn. 2018. "Implantable Neural Probes for Brain-Machine Interfaces – Current Developments and Future Prospects." *Experimental Neurobiology* 27 (6): 453–71. https://doi.org/10.5607/en.2018.27.6.453.

Collinger, Jennifer L, Brian Wodlinger, John E Downey, Wei Wang, Elizabeth C Tyler-Kabara, Douglas J Weber, Angus JC McMorland, Meel Velliste, Michael L Boninger, and Andrew B Schwartz. 2013. "7 Degree-of-Freedom Neuroprosthetic Control by an Individual with Tetraplegia." *Lancet* 381 (9866): 557–64. https://doi.org/10.1016/S0140-6736(12)61816-9.

Cooney, Ciaran, Raffaella Folli, and Damien Coyle. 2018. "Neurolinguistics Research Advancing Development of a Direct-Speech Brain-Computer Interface." *IScience* 8 (September): 103–25. https://doi.org/10.1016/j.isci.2018.09.016.

———. 2022. "Opportunities, Pitfalls and Trade-Offs in Designing Protocols for Measuring the Neural Correlates of Speech." *Neuroscience & Biobehavioral Reviews* 140 (September): 104783. https://doi.org/10.1016/j.neubiorev.2022.104783.

Cui, He, and Richard A. Andersen. 2007. "Posterior Parietal Cortex Encodes Autonomously Selected Motor Plans." *Neuron* 56 (3): 552–59. https://doi.org/10.1016/j.neuron.2007.09.031.

Dash, Debadatta, Paul Ferrari, Angel Hernandez, Daragh Heitzman, Sara G. Austin, and Jun Wang. 2020. "Neural Speech Decoding for Amyotrophic Lateral Sclerosis." In *Interspeech 2020*, 2782–86. ISCA. https://doi.org/10.21437/Interspeech.2020-3071.

Dash, Debadatta, Paul Ferrari, and Jun Wang. 2020a. "Decoding Imagined and Spoken Phrases From Non-Invasive Neural (MEG) Signals." *Frontiers in Neuroscience* 14. https://www.frontiersin.org/articles/10.3389/fnins.2020.00290.

———. 2020b. "Decoding Imagined and Spoken Phrases From Non-Invasive Neural (MEG) Signals." *Frontiers in Neuroscience* 14. https://www.frontiersin.org/article/10.3389/fnins.2020.00290.

Deschamps, Isabelle, Shari R. Baum, and Vincent L. Gracco. 2014. "On the Role of the Supramarginal Gyrus in Phonological Processing and Verbal Working Memory: Evidence from RTMS Studies." *Neuropsychologia* 53 (January): 39–46. https://doi.org/10.1016/j.neuropsychologia.2013.10.015.

Ding, Weizhong, Shian Hu, Pengju Wang, Honglei Kang, Renpeng Peng, Yimin Dong, and Feng Li. 2022. "Spinal Cord Injury: The Global Incidence, Prevalence, and Disability From the Global Burden of Disease Study 2019." *Spine* 47 (21): 1532. https://doi.org/10.1097/BRS.0000000000004417.

Feix, T., J. Romero, H. B. Schmiedmayer, A. M. Dollar, and D. Kragic. 2016. "The GRASP Taxonomy of Human Grasp Types." *IEEE Transactions on Human-Machine Systems* 46 (1): 66–77. https://doi.org/10.1109/THMS.2015.2470657.

Filik, Ruth, and Emma Barber. 2011. "Inner Speech during Silent Reading Reflects the Reader's Regional Accent." *PLoS ONE* 6 (10): e25782. https://doi.org/10.1371/journal.pone.0025782.

Filimon, Flavia, Jonathan D. Nelson, Ruey-Song Huang, and Martin I. Sereno. 2009. "Multiple Parietal Reach Regions in Humans: Cortical Representations for Visual and Proprioceptive Feedback

during on-Line Reaching." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29 (9): 2961–71. https://doi.org/10.1523/JNEUROSCI.3211-08.2009.

Flesher, Sharlene N., Jennifer L. Collinger, Stephen T. Foldes, Jeffrey M. Weiss, John E. Downey, Elizabeth C. Tyler-Kabara, Sliman J. Bensmaia, Andrew B. Schwartz, Michael L. Boninger, and Robert A. Gaunt. 2016. "Intracortical Microstimulation of Human Somatosensory Cortex." *Science Translational Medicine* 8 (361): 361ra141-361ra141. https://doi.org/10.1126/scitranslmed.aaf8083.

Gallego, Juan A., Tamar R. Makin, and Samuel D. McDougle. 2022. "Going beyond Primary Motor Cortex to Improve Brain–Computer Interfaces." *Trends in Neurosciences* 45 (3): 176–83. https://doi.org/10.1016/j.tins.2021.12.006.

Gallivan, Jason P, D Adam McLean, Kenneth F Valyear, and Jody C Culham. 2013. "Decoding the Neural Mechanisms of Human Tool Use." Edited by Dora Angelaki. *ELife* 2 (May): e00425. https://doi.org/10.7554/eLife.00425.

Garcea, Frank E., and Laurel J. Buxbaum. 2019. "Gesturing Tool Use and Tool Transport Actions Modulates Inferior Parietal Functional Connectivity with the Dorsal and Ventral Object Processing Pathways." *Human Brain Mapping* 40 (10): 2867–83. https://doi.org/10.1002/hbm.24565.

Geranmayeh, Fatemeh, Sonia L. E. Brownsett, Robert Leech, Christian F. Beckmann, Zoe Woodhead, and Richard J. S. Wise. 2012. "The Contribution of the Inferior Parietal Cortex to Spoken Language Production." *Brain and Language* 121 (1): 47–57. https://doi.org/10.1016/j.bandl.2012.02.005.

Geva, Sharon, P. Simon Jones, Jenny T. Crinion, Cathy J. Price, Jean-Claude Baron, and Elizabeth A. Warburton. 2011. "The Neural Correlates of Inner Speech Defined by Voxel-Based Lesion–Symptom Mapping." *Brain* 134 (10): 3071–82. https://doi.org/10.1093/brain/awr232.

Goldenberg, G., and J. Spatt. 2009. "The Neural Basis of Tool Use." *Brain* 132 (6): 1645–55. https://doi.org/10.1093/brain/awp080.

Goodman, James M., Gregg A. Tabot, Alex S. Lee, Aneesha K. Suresh, Alexander T. Rajan, Nicholas G. Hatsopoulos, and Sliman Bensmaia. 2019. "Postural Representations of the Hand in the Primate Sensorimotor Cortex." *Neuron* 104 (5): 1000-1009.e7. https://doi.org/10.1016/j.neuron.2019.09.004.

Guan, Charles, Tyson Aflalo, Kelly Kadlec, Jorge Gámez de Leon, Emily R. Rosario, Ausaf Bari, Nader Pouratian, and Richard A. Andersen. 2022. "Compositional Coding of Individual Finger Movements in Human Posterior Parietal Cortex and Motor Cortex Enables Ten-Finger Decoding." Preprint. Neurology. https://doi.org/10.1101/2022.12.07.22283227.

Hartwigsen, Gesa, Annette Baumgaertner, Cathy J. Price, Maria Koehnke, Stephan Ulmer, and Hartwig R. Siebner. 2010. "Phonological Decisions Require Both the Left and Right Supramarginal Gyri." *Proceedings of the National Academy of Sciences* 107 (38): 16494–99. https://doi.org/10.1073/pnas.1008121107.

Hecht, Martin, Thomas Hillemacher, Elmar Gräsel, Sebastian Tigges, Martin Winterholler, Dieter Heuss, Max-Josef Hilz, and Bernhard Neundörfer. 2002. "Subjective Experience and Coping in ALS." *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders: Official Publication of the World Federation of Neurology, Research Group on Motor Neuron Diseases* 3 (4): 225–31. https://doi.org/10.1080/146608202760839009.

Herff, Christian, Lorenz Diener, Miguel Angrick, Emily Mugler, Matthew C. Tate, Matthew A. Goldrick, Dean J. Krusienski, Marc W. Slutzky, and Tanja Schultz. 2019. "Generating Natural,

Intelligible Speech From Brain Activity in Motor, Premotor, and Inferior Frontal Cortices." *Frontiers in Neuroscience* 13. https://doi.org/10.3389/fnins.2019.01267.

Herff, Christian, Dean J. Krusienski, and Pieter Kubben. 2020. "The Potential of Stereotactic-EEG for Brain-Computer Interfaces: Current Progress and Future Directions." *Frontiers in Neuroscience* 14. https://www.frontiersin.org/article/10.3389/fnins.2020.00123.

Hochberg, Leigh R., Daniel Bacher, Beata Jarosiewicz, Nicolas Y. Masse, John D. Simeral, Joern Vogel, Sami Haddadin, et al. 2012. "Reach and Grasp by People with Tetraplegia Using a Neurally Controlled Robotic Arm." *Nature* 485 (7398): 372–75. https://doi.org/10.1038/nature11076.

Hochberg, Leigh R., Mijail D. Serruya, Gerhard M. Friehs, Jon A. Mukand, Maryam Saleh, Abraham H. Caplan, Almut Branner, David Chen, Richard D. Penn, and John P. Donoghue. 2006. "Neuronal Ensemble Control of Prosthetic Devices by a Human with Tetraplegia." *Nature* 442 (7099): 164–71. https://doi.org/10.1038/nature04970.

Hotson, Guy, David P. McMullen, Matthew S. Fifer, Matthew S. Johannes, Kapil D. Katyal, Matthew P. Para, Robert Armiger, et al. 2016. "Individual Finger Control of a Modular Prosthetic Limb Using High-Density Electrocorticography in a Human Subject." *Journal of Neural Engineering* 13 (2): 026017–026017. https://doi.org/10.1088/1741-2560/13/2/026017.

Iida, Koji, and Hiroshi Otsubo. 2017. "Stereoelectroencephalography: Indication and Efficacy." *Neurologia Medico-Chirurgica* 57 (8): 375–85. https://doi.org/10.2176/nmc.ra.2017-0008.

Ikeda, Shigeyuki, Tomohiro Shibata, Naoki Nakano, Rieko Okada, Naohiro Tsuyuguchi, Kazushi Ikeda, and Amami Kato. 2014. "Neural Decoding of Single Vowels during Covert Articulation Using Electrocorticography." *Frontiers in Human Neuroscience* 8. https://www.frontiersin.org/articles/10.3389/fnhum.2014.00125.

Indefrey, P., and W. J. M. Levelt. 2004. "The Spatial and Temporal Signatures of Word Production Components." *Cognition* 92 (1–2): 101–44. https://doi.org/10.1016/j.cognition.2002.06.001.

Jafari, Matiar, Tyson Aflalo, Srinivas Chivukula, Spencer Sterling Kellis, Michelle Armenta Salas, Sumner Lee Norman, Kelsie Pejsa, Charles Yu Liu, and Richard Alan Andersen. 2020. "The Human Primary Somatosensory Cortex Encodes Imagined Movement in the Absence of Sensory Information." *Communications Biology* 3 (1): 1–7. https://doi.org/10.1038/s42003-020-01484-1.

Johnson-Frey, Scott H. 2004a. "The Neural Bases of Complex Tool Use in Humans." *Trends in Cognitive Sciences* 8 (2): 71–78. https://doi.org/10.1016/j.tics.2003.12.002.

———. 2004b. "The Neural Bases of Complex Tool Use in Humans." *Trends in Cognitive Sciences* 8 (2): 71–78. https://doi.org/10.1016/j.tics.2003.12.002.

Kalaska, J. F. 2009. "From Intention to Action: Motor Cortex and the Control of Reaching Movements | EndNote Click." 2009. https://click.endnote.com/viewer?doi=10.1007%2F978-0-387-77064-2_8&token=WzExMjY3MSwiMTAuMTAwNy85NzgtMC0zODctNzcwNjQtMl84Il0.lpqgCpMfWNl4adX7EuIX5lyzLhA.

Kauhanen, L., T. Nykopp, J. Lehtonen, P. Jylanki, J. Heikkonen, P. Rantanen, H. Alaranta, and M. Sams. 2006. "EEG and MEG Brain-Computer Interface for Tetraplegic Patients." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14 (2): 190–93. https://doi.org/10.1109/TNSRE.2006.875546.

Kellis, Spencer, Kai Miller, Kyle Thomson, Richard Brown, Paul House, and Bradley Greger. 2010. "Decoding Spoken Words Using Local Field Potentials Recorded from the Cortical Surface."

Journal of Neural Engineering 7 (5): 056007. https://doi.org/10.1088/1741-2560/7/5/056007.

Klaes, Christian, Spencer Kellis, Tyson Aflalo, Brian Lee, Kelsie Pejsa, Kathleen Shanfield, Stephanie Hayes-Jackson, et al. 2015. "Hand Shape Representations in the Human Posterior Parietal Cortex." *The Journal of Neuroscience* 35 (46): 15466–76. https://doi.org/10.1523/JNEUROSCI.2747-15.2015.

Kobak, Dmitry, Wieland Brendel, Christos Constantinidis, Claudia E Feierstein, Adam Kepecs, Zachary F Mainen, Xue-Lian Qi, Ranulfo Romo, Naoshige Uchida, and Christian K Machens. 2016. "Demixed Principal Component Analysis of Neural Population Data." Edited by Mark CW van Rossum. *ELife* 5 (April): e10989. https://doi.org/10.7554/eLife.10989.

Langland-Hassan, Peter, and Agustin Vicente. 2018. *Inner Speech: New Voices*. Oxford University Press.

Leuthardt, Eric, Xiao-mei Pei, Jonathan Breshears, Charles Gaona, Mohit Sharma, Zachary Freudenburg, Dennis Barbour, and Gerwin Schalk. 2012. "Temporal Evolution of Gamma Activity in Human Cortex during an Overt and Covert Word Repetition Task." *Frontiers in Human Neuroscience* 6. https://doi.org/10.3389/fnhum.2012.00099.

Lœvenbruck, Hélène, Monica Baciu, Christoph Segebarth, and Christian Abry. 2005. "The Left Inferior Frontal Gyrus under Focus: An FMRI Study of the Production of Deixis via Syntactic Extraction and Prosodic Focus." *Journal of Neurolinguistics* 18 (3): 237–58. https://doi.org/10.1016/j.jneuroling.2004.12.002.

Lopez-Bernal, Diego, David Balderas, Pedro Ponce, and Arturo Molina. 2022. "A State-of-the-Art Review of EEG-Based Imagined Speech Decoding." *Frontiers in Human Neuroscience* 16 (April): 867281. https://doi.org/10.3389/fnhum.2022.867281.

Luo, Shiyu, Qinwan Rabbani, and Nathan E. Crone. 2022. "Brain-Computer Interface: Applications to Speech Decoding and Synthesis to Augment Communication." *Neurotherapeutics: The Journal of the American Society for Experimental NeuroTherapeutics*, January. https://doi.org/10.1007/s13311-022-01190-2.

Makin, Joseph G., David A. Moses, and Edward F. Chang. 2020. "Machine Translation of Cortical Activity to Text with an Encoder–Decoder Framework." *Nature Neuroscience* 23 (4): 575–82. https://doi.org/10.1038/s41593-020-0608-8.

Makkonen, Tanja, Hanna Ruottinen, Riitta Puhto, Mika Helminen, and Johanna Palmio. 2018. "Speech Deterioration in Amyotrophic Lateral Sclerosis (ALS) after Manifestation of Bulbar Symptoms." *International Journal of Language & Communication Disorders* 53 (2): 385–92. https://doi.org/10.1111/1460-6984.12357.

Malmivuo, Jaakko. 2012. "Comparison of the Properties of EEG and MEG in Detecting the Electric Activity of the Brain." *Brain Topography* 25 (1): 1–19. https://doi.org/10.1007/s10548-011-0202-1.

Martin, Stephanie, Peter Brunner, Iñaki Iturrate, José del R. Millán, Gerwin Schalk, Robert T. Knight, and Brian N. Pasley. 2016. "Word Pair Classification during Imagined Speech Using Direct Brain Recordings." *Scientific Reports* 6 (1): 25803. https://doi.org/10.1038/srep25803.

Martin, Stephanie, Iñaki Iturrate, José del R. Millán, Robert T. Knight, and Brian N. Pasley. 2018. "Decoding Inner Speech Using Electrocorticography: Progress and Challenges Toward a Speech Prosthesis." *Frontiers in Neuroscience* 12. https://doi.org/10.3389/fnins.2018.00422.

McDowell, Tomás, Nicholas P. Holmes, Alan Sunderland, and Martin Schürmann. 2018. "TMS over the Supramarginal Gyrus Delays Selection of Appropriate Grasp Orientation during

Reaching and Grasping Tools for Use." *Cortex* 103 (June): 117–29. https://doi.org/10.1016/j.cortex.2018.03.002.

Mehta, Paul. 2018. "Prevalence of Amyotrophic Lateral Sclerosis — United States, 2015." *MMWR. Morbidity and Mortality Weekly Report* 67. https://doi.org/10.15585/mmwr.mm6746a1.

Metzger, Sean L., Jessie R. Liu, David A. Moses, Maximilian E. Dougherty, Margaret P. Seaton, Kaylo T. Littlejohn, Josh Chartier, et al. 2022. "Generalizable Spelling Using a Speech Neuroprosthesis in an Individual with Severe Limb and Vocal Paralysis." *Nature Communications* 13 (1): 6510. https://doi.org/10.1038/s41467-022-33611-3.

Michaels, Jonathan, and Hans Scherberger. 2017. *Population Coding of Grasp and Laterality-Related Information in the Macaque Fronto-Parietal Network*. https://doi.org/10.1101/179184.

Moses, David A., Sean L. Metzger, Jessie R. Liu, Gopala K. Anumanchipalli, Joseph G. Makin, Pengfei F. Sun, Josh Chartier, et al. 2021. "Neuroprosthesis for Decoding Speech in a Paralyzed Person with Anarthria." *New England Journal of Medicine* 385 (3): 217–27. https://doi.org/10.1056/NEJMoa2027540.

Mullin, Emily. 2022. "This Man Set the Record for Wearing a Brain-Computer Interface." *Wired*, 2022. https://www.wired.com/story/this-man-set-the-record-for-wearing-a-brain-computer-interface/.

Murata, A., V. Gallese, G. Luppino, M. Kaseda, and H. Sakata. 2000. "Selectivity for the Shape, Size, and Orientation of Objects for Grasping in Neurons of Monkey Parietal Area AIP." *Journal of Neurophysiology* 83 (5): 2580–2601. https://doi.org/10.1152/jn.2000.83.5.2580.

Murata, Akira, Luciano Fadiga, Leonardo Fogassi, Vittorio Gallese, Vassilis Raos, and Giacomo Rizzolatti. 1997. "Object Representation in the Ventral Premotor Cortex (Area F5) of the Monkey." *Journal of Neurophysiology* 78 (4): 2226–30. https://doi.org/10.1152/jn.1997.78.4.2226.

Muret, Dollyane, Victoria Root, Paulina Kieliba, Danielle Clode, and Tamar R. Makin. 2022. "Beyond Body Maps: Information Content of Specific Body Parts Is Distributed across the Somatosensory Homunculus." *Cell Reports* 38 (11): 110523. https://doi.org/10.1016/j.celrep.2022.110523.

National Spinal Cord Injury Statistical Center. 2022. "Traumatic Spinal Cord Injury Facts and Figures at a Glance."

Nicolas-Alonso, Luis Fernando, and Jaime Gomez-Gil. 2012. "Brain Computer Interfaces, a Review." *Sensors (Basel, Switzerland)* 12 (2): 1211–79. https://doi.org/10.3390/s120201211.

Norris, Forbes, Rodger Shepherd, Eric Denys, Kwei U, Eichiro Mukai, Linda Elias, Dolores Holden, and Holten Norris. 1993. "Onset, Natural History and Outcome in Idiopathic Adult Motor Neuron Disease." *Journal of the Neurological Sciences* 118 (1): 48–55. https://doi.org/10.1016/0022-510X(93)90245-T.

Oberhuber, M., T. M. H. Hope, M. L. Seghier, O. Parker Jones, S. Prejawa, D. W. Green, and C. J Price. 2016. "Four Functionally Distinct Regions in the Left Supramarginal Gyrus Support Word Processing." *Cerebral Cortex (New York, NY)* 26 (11): 4212–26. https://doi.org/10.1093/cercor/bhw251.

Okorokova, Elizaveta V., James M. Goodman, Nicholas G. Hatsopoulos, and Sliman J. Bensmaia. 2020. "Decoding Hand Kinematics from Population Responses in Sensorimotor Cortex during Grasping." *Journal of Neural Engineering* 17 (4): 046035. https://doi.org/10.1088/1741-2552/ab95ea.

Orban, Guy A., and Fausto Caruana. 2014. "The Neural Basis of Human Tool Use." *Frontiers in Psychology* 5 (April). https://doi.org/10.3389/fpsyg.2014.00310.

Osiurak, François, and Arnaud Badets. 2016. “Tool Use and Affordance: Manipulation-Based versus Reasoning-Based Approaches.” *Psychological Review* 123 (5): 534–68. https://doi.org/10.1037/rev0000027.

Peeters, R., L. Simone, K. Nelissen, M. Fabbri-Destro, W. Vanduffel, G. Rizzolatti, and G. A. Orban. 2009. “The Representation of Tool Use in Humans and Monkeys: Common and Uniquely Human Features.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29 (37): 11523–39. https://doi.org/10.1523/JNEUROSCI.2040-09.2009.

Peeters, Ronald R., Giacomo Rizzolatti, and Guy A. Orban. 2013. “Functional Properties of the Left Parietal Tool Use Region.” *NeuroImage* 78 (September): 83–93. https://doi.org/10.1016/j.neuroimage.2013.04.023.

Pei, Xiaomei, Dennis L. Barbour, Eric C. Leuthardt, and Gerwin Schalk. 2011. “Decoding Vowels and Consonants in Spoken and Imagined Words Using Electrocorticographic Signals in Humans.” *Journal of Neural Engineering* 8 (4): 046028. https://doi.org/10.1088/1741-2560/8/4/046028.

Pei, Xiaomei, Eric C. Leuthardt, Charles M. Gaona, Peter Brunner, Jonathan R. Wolpaw, and Gerwin Schalk. 2011a. “Spatiotemporal Dynamics of Electrocorticographic High Gamma Activity during Overt and Covert Word Repetition.” *NeuroImage* 54 (4): 2960–72. https://doi.org/10.1016/j.neuroimage.2010.10.029.

———. 2011b. “Spatiotemporal Dynamics of Electrocorticographic High Gamma Activity during Overt and Covert Word Repetition.” *NeuroImage* 54 (4): 2960–72. https://doi.org/10.1016/j.neuroimage.2010.10.029.

Perini, Francesca, Thomas Powell, Simon J. Watt, and Paul E. Downing. 2020. “Neural Representations of Haptic Object Size in the Human Brain Revealed by Multivoxel FMRI Patterns.” *Journal of Neurophysiology* 124 (1): 218–31. https://doi.org/10.1152/jn.00160.2020.

Perrone-Bertolotti, M., L. Rapin, J. -P. Lachaux, M. Baciu, and H. Lœvenbruck. 2014. “What Is That Little Voice inside My Head? Inner Speech Phenomenology, Its Role in Cognitive Performance, and Its Relation to Self-Monitoring.” *Behavioural Brain Research* 261 (March): 220–39. https://doi.org/10.1016/j.bbr.2013.12.034.

Potok, Weronika, Adam Maskiewicz, Gregory Króliczak, and Mattia Marangon. 2019. “The Temporal Involvement of the Left Supramarginal Gyrus in Planning Functional Grasps: A Neuronavigated TMS Study.” *Cortex* 111 (February): 16–34. https://doi.org/10.1016/j.cortex.2018.10.010.

Pratt, Ethan J., Micah Ledbetter, Ricardo Jiménez-Martínez, Benjamin Shapiro, Amelia Solon, Geoffrey Z. Iwata, Steve Garber, et al. 2021. “Kernel Flux: Optical and Quantum Sensing and Precision Metrology 2021.” Edited by Selim M. Shahriar and Jacob Scheuer. *Optical and Quantum Sensing and Precision Metrology*, Proceedings of SPIE - The International Society for Optical Engineering, . https://doi.org/10.1117/12.2581794.

Price, C. J., C. J. Moore, G. W. Humphreys, and R. J. S. Wise. 1997. “Segregating Semantic from Phonological Processes during Reading.” *Journal of Cognitive Neuroscience* 9 (6): 727–33. https://doi.org/10.1162/jocn.1997.9.6.727.

Price, Cathy J. 2010. “The Anatomy of Language: A Review of 100 FMRI Studies Published in 2009.” *Annals of the New York Academy of Sciences* 1191 (1): 62–88. https://doi.org/10.1111/j.1749-6632.2010.05444.x.

Proix, Timothée, Jaime Delgado Saa, Andy Christen, Stephanie Martin, Brian N. Pasley, Robert T. Knight, Xing Tian, et al. 2022. “Imagined Speech Can Be Decoded from Low- and Cross-

Frequency Intracranial EEG Features." *Nature Communications* 13 (1): 48. https://doi.org/10.1038/s41467-021-27725-3.

Rabbani, Qinwan, Griffin Milsap, and Nathan E. Crone. 2019. "The Potential for a Speech Brain–Computer Interface Using Chronic Electrocorticography." *Neurotherapeutics* 16 (1): 144–65. https://doi.org/10.1007/s13311-018-00692-2.

Reynaud, Emanuelle, Jordan Navarro, Mathieu Lesourd, and François Osiurak. 2019. "To Watch Is to Work: A Review of NeuroImaging Data on Tool Use Observation Network." *Neuropsychology Review* 29 (4): 484–97. https://doi.org/10.1007/s11065-019-09418-3.

Rizzolatti, G., G. Luppino, and M. Matelli. 1998. "The Organization of the Cortical Motor System: New Concepts." *Electroencephalography and Clinical Neurophysiology* 106 (4): 283–96.

Rosenthal, Isabelle A., Luke Bashford, Spencer Kellis, Kelsie Pejsa, Brian Lee, Charles Liu, and Richard A. Andersen. 2022. "S1 Represents Multisensory Contexts and Somatotopic Locations within and Outside the Bounds of the Cortical Homunculus." bioRxiv. https://doi.org/10.1101/2022.08.29.505313.

Roussel, Philémon, Gaël Le Godais, Florent Bocquelet, Marie Palma, Jiang Hongjie, Shaomin Zhang, Anne-Lise Giraud, et al. 2020. "Observation and Assessment of Acoustic Contamination of Electrophysiological Brain Signals during Speech Production and Sound Perception." *Journal of Neural Engineering* 17 (5): 056028. https://doi.org/10.1088/1741-2552/abb25e.

Roux, Franck-Emmanuel, Jean-Baptiste Durand, Mélanie Jucla, Emilie Réhault, Marion Reddy, and Jean-François Démonet. 2012. "Segregation of Lexical and Sub-Lexical Reading Processes in the Left Perisylvian Cortex." *PLoS ONE* 7 (11): e50665. https://doi.org/10.1371/journal.pone.0050665.

Rutishauser, Ueli, Tyson Aflalo, Emily R. Rosario, Nader Pouratian, and Richard A. Andersen. 2018. "Single-Neuron Representation of Memory Strength and Recognition Confidence in Left Human Posterior Parietal Cortex." *Neuron* 97 (1): 209-220.e3. https://doi.org/10.1016/j.neuron.2017.11.029.

Sakata, H. 1995. "Neural Mechanisms of Visual Guidance of Hand Action in the Parietal Cortex of the Monkey | Cerebral Cortex | Oxford Academic." 1995. https://academic.oup.com/cercor/article-abstract/5/5/429/375668?redirectedFrom=fulltext.

Schaffelhofer, Stefan, Andres Agudelo-Toro, and Hansjörg Scherberger. 2015a. "Decoding a Wide Range of Hand Configurations from Macaque Motor, Premotor, and Parietal Cortices." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 35 (3): 1068–81. https://doi.org/10.1523/JNEUROSCI.3594-14.2015.

———. 2015b. "Decoding a Wide Range of Hand Configurations from Macaque Motor, Premotor, and Parietal Cortices." *Journal of Neuroscience* 35 (3): 1068–81. https://doi.org/10.1523/JNEUROSCI.3594-14.2015.

Schaffelhofer, Stefan, and Hansjörg Scherberger. 2016. "Object Vision to Hand Action in Macaque Parietal, Premotor, and Motor Cortices." *ELife* 5. https://doi.org/10.7554/eLife.15278.

Seghier, Mohamed L., François Lazeyras, Alan J. Pegna, Jean-Marie Annoni, Ivan Zimine, Eugène Mayer, Christoph M. Michel, and Asaid Khateb. 2004. "Variability of FMRI Activation during a Phonological and Semantic Language Task in Healthy Subjects." *Human Brain Mapping* 23 (3): 140–55. https://doi.org/10.1002/hbm.20053.

Shokoueinejad, Mehdi, Dong-Wook Park, Yei Hwan Jung, Sarah K. Brodnick, Joseph Novello, Aaron Dingle, Kyle I. Swanson, et al. 2019. "Progress in the Field of Micro-Electrocorticography." *Micromachines* 10 (1): 62. https://doi.org/10.3390/mi10010062.

Simeral, J D, S-P Kim, M J Black, J P Donoghue, and L R Hochberg. 2011. “Neural Control of Cursor Trajectory and Click by a Human with Tetraplegia 1000 Days after Implant of an Intracortical Microelectrode Array.” *Journal of Neural Engineering* 8 (2): 025027. https://doi.org/10.1088/1741-2560/8/2/025027.

Sliwinska, Magdalena Wiktoria Wiktoria, Manali Khadilkar, Jonathon Campbell-Ratcliffe, Frances Quevenco, and Joseph T. Devlin. 2012a. “Early and Sustained Supramarginal Gyrus Contributions to Phonological Processing.” *Frontiers in Psychology* 3. https://doi.org/10.3389/fpsyg.2012.00161.

———. 2012b. “Early and Sustained Supramarginal Gyrus Contributions to Phonological Processing.” *Frontiers in Psychology* 3. https://doi.org/10.3389/fpsyg.2012.00161.

Smith, Eimear, and Mark Delargy. 2005. “Locked-in Syndrome.” *BMJ* 330 (7488): 406–9. https://doi.org/10.1136/bmj.330.7488.406.

Snoek, G. J., M. J. IJzerman, H. J. Hermens, D. Maxwell, and F. Biering-Sorensen. 2004. “Survey of the Needs of Patients with Spinal Cord Injury: Impact and Priority for Improvement in Hand Function in Tetraplegics.” *Spinal Cord* 42 (9): 526–32. https://doi.org/10.1038/sj.sc.3101638.

Soroush, Pedram Z., Christian Herff, Stephanie K. Ries, Jerry J. Shih, Tanja Schultz, and Dean J. Krusienski. 2022. “The Nested Hierarchy of Overt, Mouthed, and Imagined Speech Activity Evident in Intracranial Recordings.” bioRxiv. https://doi.org/10.1101/2022.08.04.502829.

Stavisky, Sergey D, Francis R Willett, Guy H Wilson, Brian A Murphy, Paymon Rezaii, Donald T Avansino, William D Memberg, et al. 2019. “Neural Ensemble Dynamics in Dorsal Motor Cortex during Speech in People with Paralysis.” Edited by Tamar R Makin, Barbara G Shinn-Cunningham, Tamar R Makin, Juan Álvaro Gallego, and Sophie K Scott. *ELife* 8 (December): e46015. https://doi.org/10.7554/eLife.46015.

Stoeckel, Cornelia, Patricia M. Gough, Kate E. Watkins, and Joseph T. Devlin. 2009. “Supramarginal Gyrus Involvement in Visual Word Recognition.” *Cortex*, Special Issue on “The Contribution of TMS to Structure-Function Mapping in the Human Brain. Action, Perception and Higher Functions,” 45 (9): 1091–96. https://doi.org/10.1016/j.cortex.2008.12.004.

Styrkowiec, Piotr P., Agnieszka M. Nowik, and Gregory Króliczak. 2019. “The Neural Underpinnings of Haptically Guided Functional Grasping of Tools: An FMRI Study.” *NeuroImage* 194 (July): 149–62. https://doi.org/10.1016/j.neuroimage.2019.03.043.

Taira. 1998. “Parietal Cortex Neurons of the Monkey Related to the Visual Guidance of Hand Movement | SpringerLink.” 1998. https://link.springer.com/article/10.1007/BF00232190.

Townsend, Benjamin R., Erk Subasi, and Hansjörg Scherberger. 2011. “Grasp Movement Decoding from Premotor and Parietal Cortex.” *Journal of Neuroscience* 31 (40): 14386–98. https://doi.org/10.1523/JNEUROSCI.2451-11.2011.

Vingerhoets, Guy. 2014. “Contribution of the Posterior Parietal Cortex in Reaching, Grasping, and Using Objects and Tools.” *Frontiers in Psychology* 5 (March). https://doi.org/10.3389/fpsyg.2014.00151.

Wandelt, Sarah K., Spencer Kellis, David A. Bjånes, Kelsie Pejsa, Brian Lee, Charles Liu, and Richard A. Andersen. 2022a. “Decoding Grasp and Speech Signals from the Cortical Grasp Circuit in a Tetraplegic Human.” *Neuron*, March. https://doi.org/10.1016/j.neuron.2022.03.009.

———. 2022b. “Decoding Grasp and Speech Signals from the Cortical Grasp Circuit in a Tetraplegic Human.” *Neuron* 110 (11): 1777-1787.e3. https://doi.org/10.1016/j.neuron.2022.03.009.

Wang, Meijian, and Liang Guo. 2020. “Intracortical Electrodes.” In *Neural Interface Engineering: Linking the Physical World and the Nervous System*, edited by Liang Guo, 67–94. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-41854-0_4.

Wang, Meng, Guangye Li, Shize Jiang, Zixuan Wei, Jie Hu, Liang Chen, and Dingguo Zhang. 2020. “Enhancing Gesture Decoding Performance Using Signals from Posterior Parietal Cortex: A Stereo-Electroencephalograhy (SEEG) Study.” *Journal of Neural Engineering* 17 (4): 046043. https://doi.org/10.1088/1741-2552/ab9987.

Willett, Francis, Erin Kunz, Chaofei Fan, Donald Avansino, Guy Wilson, Eun Young Choi, Foram Kamdar, et al. 2023. “A High-Performance Speech Neuroprosthesis.” bioRxiv. https://doi.org/10.1101/2023.01.21.524489.

Willett, Francis R., Donald T. Avansino, Leigh R. Hochberg, Jaimie M. Henderson, and Krishna V. Shenoy. 2021. “High-Performance Brain-to-Text Communication via Handwriting.” *Nature* 593 (7858): 249–54. https://doi.org/10.1038/s41586-021-03506-2.

Wilson, Guy H., Sergey D. Stavisky, Francis R. Willett, Donald T. Avansino, Jessica N. Kelemen, Leigh R. Hochberg, Jaimie M. Henderson, Shaul Druckmann, and Krishna V. Shenoy. 2020. “Decoding Spoken English from Intracortical Electrode Arrays in Dorsal Precentral Gyrus.” *Journal of Neural Engineering* 17 (6): 066007. https://doi.org/10.1088/1741-2552/abbfef.

Zhang, Carey Y., Tyson Aflalo, Boris Revechkis, Emily R. Rosario, Debra Ouellette, Nader Pouratian, and Richard A. Andersen. 2017. “Partially Mixed Selectivity in Human Posterior Parietal Association Cortex.” *Neuron* 95 (3): 697-708.e4. https://doi.org/10.1016/j.neuron.2017.06.040.