Simultaneous Real-Time Monitoring of Multiple Cortical Systems

Journal:	*Journal of Neural Engineering*
Manuscript ID:	JNE-100098.R2
Manuscript Type:	Paper
Date Submitted by the Author:	02-Jun-2014
Complete List of Authors:	Gupta, Disha; Burke-Cornell Medical Research Institute, Early Brain Injury and Motor Recovery Hill, Nicholas; Wadsworth Center, Brain Computer Interfacing Brunner, Peter; Albany Medical College, Dept. of Neurology Gunduz, Aysegul; Albany Medical College, Dept. of Neurology Ritaccio, Anthony; Albany Medical College, Dept. of Neurology Schalk, Gerwin; Wadsworth Center, Brain Computer Interfacing
Article Keywords:	Electrocorticography, Movement Intention, Auditory Processing, Realtime Decoding, Simultaneous Decoding
Abstract:	Objective: Real-time monitoring of the brain is potentially valuable for performance monitoring, communication, training or rehabilitation. In natural situations, the brain performs a complex mix of various sensory, motor, or cognitive functions. Thus, real-time brain monitoring would be most valuable if (a) it could decode information from multiple brain systems simultaneously, and (b) this decoding of each brain system were robust to variations in the activity of other (unrelated) brain systems. Previous studies showed that it is possible to decode some information from di fferent brain systems in retrospect and/or in isolation. In our study, we set out to determine whether it is possible to simultaneously decode important information about a user from diff erent brain systems in real time, and to evaluate the impact of concurrent activity in diff erent brain systems on decoding performance. Approach: We study these questions using electrocorticographic (ECoG) signals recorded in humans. We first document procedures for generating stable decoding models given little training data, and then report their use for onl⍥ ine and for realtime decoding from 12 subjects (6 for onl⍥ ine parameter optimization, 6 for online experimentation). The subjects engage in tasks that involve movement intention, movement execution and auditory functions, separately, and then simultaneously. Main results: Our real-time results demonstrate that our system can identify intention and movement periods in single trials with an accuracy of 80.4% and 86.8%, respectively (where 50% would be expected by chance). Simultaneously, the decoding of the power envelope of an auditory stimulus resulted in an average correlation coefficient of 0.37 between the actual and decoded power envelope. These decoders were trained separately and executed simultaneously in real time.

	Significance: This study yielded the first demonstration that it is possible to decode simultaneously the functional activity of multiple independent brain systems. Our comparison of univariate and multivariate decoding strategies, and our analysis of the influence of their decoding parameters, provides benchmarks and guidelines for future research on this topic.

Simultaneous Real-Time Monitoring of Multiple Cortical Systems

Disha Gupta[1,2,3], N. Jeremy Hill[1,6], Peter Brunner[1,2], Aysegul Gunduz[1,2,4], Anthony L. Ritaccio[2], Gerwin Schalk[1,2,5]

[1] Wadsworth Center, New York State Department of Health, Albany, NY

[2] Department of Neurology, Albany Medical College, Albany, NY

[3] Early Brain Injury Recovery Program, Burke-Cornell Medical Research Institute, White Plains, NY

[4] J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL

[5] Department of Biomedical Sciences, State University of New York at Albany

[6] Translational Neurological Research Laboratory, Helen Hayes Hospital, West Haverstraw, NY

[*] Correspondence: Gerwin Schalk, Wadsworth Center, C650 Empire State Plaza, Albany, NY 12201, USA. e-mail: schalk@wadsworth.org

Abstract.

Objective Real-time monitoring of the brain is potentially valuable for performance monitoring, communication, training or rehabilitation. In natural situations, the brain performs a complex mix of various sensory, motor, or cognitive functions. Thus, real-time brain monitoring would be most valuable if (a) it could decode information from multiple brain systems simultaneously, and (b) this decoding of each brain system were robust to variations in the activity of other (unrelated) brain systems. Previous studies showed that it is possible to decode some information from different brain systems in retrospect and/or in isolation. In our study, we set out to determine whether it is possible to simultaneously decode important information about a user from different brain systems in real time, and to evaluate the impact of concurrent activity in different brain systems on decoding performance.

Approach We study these questions using electrocorticographic (ECoG) signals recorded in humans. We first document procedures for generating stable decoding models given little training data, and then report their use for offline and for real-time decoding from 12 subjects (6 for offline parameter optimization, 6 for online experimentation). The subjects engage in tasks that involve movement intention, movement execution and auditory functions, separately, and then simultaneously.

Main results Our real-time results demonstrate that our system can identify intention and movement periods in single trials with an accuracy of 80.4% and 86.8%, respectively (where 50% would be expected by chance). Simultaneously, the decoding of the power envelope of an auditory stimulus resulted in an average correlation coefficient of 0.37 between the actual and decoded power envelope. These decoders were trained separately and executed simultaneously in real time.

Significance This study yielded the first demonstration that it is possible to decode simultaneously the functional activity of multiple independent brain systems. Our comparison of univariate and multivariate decoding strategies, and our analysis of the influence of their decoding parameters, provides benchmarks and guidelines for future

research on this topic.

1. Introduction

Researchers in the field of neural engineering envision that real-time decoding of brain activity can be used for performance monitoring, functional restoration and/or rehabilitation [1, 2, 3, 4, 5, 6, 7, 8, 9]. One critical and typically ignored issue of such brain-based applications is that real-world tasks usually directly or indirectly involve multiple sensory, motor and/or cognitive brain processes. Thus, decoding information from the brain may need to access information from different brain systems and needs to be robust to variations in other (unrelated) brain processes. Previous (mostly offline) studies have shown that it is possible to relate specific functional parameters to brain signal features for isolated brain functions. These demonstrations usually involved the motor or auditory systems, and were accomplished using non-invasive electroencephalography (EEG) [10, 11, 3], invasive electrocorticography (ECoG) [12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28], or invasive single-neuron recordings [29, 30, 31, 32, 33, 34]. However, it remains unclear whether the relationships between different functional parameters and brain signals are conserved during real-world tasks, which tend to be complex and multi-modal and often occur in varied and unpredictable contexts. We aim to determine whether it is possible to decode multiple brain systems when they are engaged simultaneously. In doing so, we also set out to assess the effect of various decoding parameters on decoding, and the robustness of each task to the activity of the other brain processes. If they could be further validated in other contexts, these procedures would have important implications for the practical utility of brain-based systems for monitoring, functional restoration or rehabilitation. In the present study, we simultaneously extract from ECoG signals, in real time, parameters relating to motor preparation, motor execution, and auditory processing. Thus, we provide the first demonstration of simultaneous multiple-system decoding, as well as the first evidence that the decoding of an individual brain system can be stable even when other brain processes are engaged.

In our study, we use ECoG signals from subdural grids implanted on the surface of the brain. In contrast to metabolic methods such as functional magnetic resonance imaging (fMRI), ECoG has high temporal resolution (< 1 msec). In contrast to single-unit recordings, ECoG signals provide detailed information about brain function across large areas of the brain, and also appear to have distinct advantages in signal robustness [35]. In contrast to non-invasive EEG signals, ECoG has much greater signal-to-noise ratio [20], minimal artifacts, higher spatial resolution (< 1 cm), broader bandwidth (0–500 Hz) and higher amplitude (50–100 μV for ECoG in contrast to 10–20 μV for EEG). ECoG signals in humans are usually acquired from people with epilepsy who undergo invasive pre-surgical monitoring for localization of epileptic foci. Recent offline studies have shown that ECoG amplitudes in certain frequency bands carry substantial task-related information, such as motor execution and planning, auditory processing and visual-spatial attention [36, 37, 17, 23, 38, 16, 22, 18, 39, 27]. Most of this information is captured in the high gamma range (around 70–110 Hz), suggesting that high gamma

activity provides reliable information about neural activity of local cortical populations. While these properties of ECoG are attractive, the present clinical circumstances of data collection add several situation-based challenges and limitations, in particular for real-time research studies. Efficient and effective ECoG-based research requires robust and general protocols [40]. Data acquisition and implementation must be carried out within a short period of 2–3 days, constrained by post-surgery recovery, clinical monitoring and mapping procedures, visitors, epileptic seizures, and variations in the subject's health, alertness and motivation. Thus, and probably not unlike eventual real-world application of such technologies, we must work within restricted time and space to set up and adapt data acquisition and processing systems. These systems must be portable, interruptible, compatible with the clinical environment required for a post-surgical patient, and resistant to the signal interference generated by medical equipment in the patient's room.

These requirements are even more severe for the multi-modal real-time decoding application discussed here. The processing pipelines for offline analysis and for real-time decoding should be identical, so that results from offline analysis can be transferred seamlessly to the real-time setting. In practice, this means that both real-time and offline pipelines must work within the constraints imposed by both the real-time nature of the application and the clinical setting of ECoG measurement. This means that (i) there is limited time for pre-processing, artifact detection and removal, and adaptation of the many decoder parameters to the characteristics of the subject and of the environment (this is a particular problem if the decoder is not robust to changes in these characteristics from session to session, requiring time-consuming repetition of parameter optimization); (ii) the total amount of data is limited, and within this, calibration data must strictly precede test data, and the amount of calibration data must be minimized to maximize the amount of test data for validation of decoder performance; (iii) signal processing and decoding algorithms, as well as visualization routines, must be simple enough, and efficiently implemented, to run in real-time on the available computing hardware; (iv) the hardware components must be assembled into a self-contained, mobile system that can be moved around a crowded hospital environment and removed quickly in an emergency; and (v) the software must allow measurement to be interrupted without loss of data, resumed quickly, and repeated efficiently.

We worked through these challenges and constraints for conducting the four phases of our study. These phases included: method optimization, generalization testing, calibration, and real-time testing. For the first two phases, we used ECoG data collected during motor preparation, motor movements, and auditory processing from an initial group of 6 subjects. These first two phases resulted in optimized procedures for data collection, artifact rejection, signal pre-processing, and classifier training. We then applied these optimized procedures to 6 new subjects. In 5 of these subjects, we realized real-time experiments that simultaneously decoded aspects of movement planning, movement execution and auditory processing.

2. Methods and Materials

All subjects were epilepsy patients that were undergoing invasive ECoG monitoring (see Section 2.1 for details). Our study contained four phases. Phases I and II involved purely offline (i.e. retrospective) analyses of data from the first six subjects. Phases III–IV were each performed as part of the real-time study with six different subjects. An overview of each of these phases is given in the following paragraphs:

Phase I (offline optimization): We first retrospectively investigated the effect of various decoding approaches and parameter settings using the data from the first measurement session performed by subjects 1–6. ECoG data were recorded as described in Section 2.3, together with stimulus timing information, while the subjects performed the experimental tasks detailed in Section 2.2 in separate runs. We aimed to acquire at least 300 trials of an 8-target center-out joystick task, and two runs, each 3–5 minutes in length, of a music listening task. In addition, Brodmann-area assignments (see Section 2.4) and functional maps (Section 2.5) were obtained for each subject. These data formed a basis for offline exploration, by cross-validation, of the decoding parameters and strategies for optimal real-time performance for each of these tasks. This is described in Section 2.6. Three decoders were optimized: one for the intention to move (motor planning in the joystick task), one for movement (motor execution in the joystick task), and one for auditory processing (from the listening task). Optimization results are presented in Sections 3.1.

Phase II (generalization): For 4 of the first 6 subjects, we had the opportunity to repeat the data collection session on a subsequent day. For these datasets, the optimal decoders from phase I were used to assess how well our system could generalize and transfer between sessions that were widely separated in time.

Phase III (calibration and decoding evaluation on separate tasks): For each of the subjects 7–12, we acquired calibration data as well as Brodmann-area assignments and functional mapping information, as described above for phase I. As for subjects 1–6, the different tasks were recorded in separate runs. Decoders for intention, movement and auditory processing were then configured separately based on these data, using the parameter settings that had been found to be optimal in Phase I. Once the data had been recorded, this process of configuration lasted no more than a few minutes. A preview of the decoders' performance was estimated by re-training them on only the first 80% of the data, and evaluating them on the unseen final 20%. Real-time performance was also subsequently captured for each of these separate tasks.

Phase IV (simultaneous multi-modal decoding): For each of the subjects 7, 8, 9, 10 and 12, we acquired data during a subsequent multimodal real-time session. Since there was no coverage of motor or premotor cortex in subject 11, we assessed only auditory processing in this subject. The other subjects performed the center-out joystick task in the presence of background music. Thus, for the first time, decoders trained on separate tasks could be assessed during multi-modal cortical processing. During

Table 1. Clinical profiles of the subjects who participated in the study

Subject	Age	Sex	Handedness	Performance IQ	Epilepsy Classification	Num. of Electrodes
1	29	F	R	136	Left temporal	96
2	56	M	R	87	Left temporal	101
3	25	M	R	not known	Left frontal	96
4	25	M	R	114	Right frontal	100
5	26	M	R	100	Right temporal	111
6	45	M	R	95	Left temporal	58
7	49	F	L	99	Left temporal	69
8	52	M	L	91	Left parietal	64
9	29	F	R	95	Left temporal	120
10	45	F	L	84	Left temporal	61
11	60	M	R	75	Left parieto-occipital	59
12	26	F	R	not known	Left temporal	128

this session, the experimenter's control screen allowed visualization of the functional activations from the three decoders (intention, movement and auditory processing) in real time. Decoder outputs were displayed in real time in the form of bar gauges, together with scrolling graphs that captured the last 10 seconds' history of the time-varying signals. To deal with drifts in signal offset as a function of time, the gauges continuously rescaled themselves such that the vertical axis contained everything between the 1^{st} and 99^{th} percentiles of the values output over the preceding 10 seconds. Elements of the real-time visualization screen are shown in Figure 2. The real-time sessions were also captured as videos, with subjects' consent. The video recorded the screen view as presented to the subject, synchronized in time with the experimenter's control screen view. Decoding performance is given in Section 3.5. Movie S1 in the supplementary material shows an excerpt from the video recording of subject 12.

2.1. Human Subjects

The ECoG data in this study were collected at the Epilepsy Monitoring Unit at Albany Medical Center (AMC) from twelve subjects who were candidates for resection surgery to treat drug-resistant epilepsy. They underwent invasive evaluation for surgical planning, where an array of electrodes was temporarily placed on the surface of the brain for the purposes of localizing the seizure focus and delineating eloquent cortex. As described in [40], clinical brain signal monitoring and review was not compromised at any time as we used a connector that split the cables coming from the patient's implanted electrode array into one set that was connected to the clinical monitoring system and another that acquired data for research purposes. All data were collected with approval by the Institutional Review Board of AMC, following subjects' informed consent and contingent on their clinical state, and willingness at the time of measurement. A brief clinical profile of the subjects is given in Table 1, and the spatial arrangement of their electrodes is shown in Figure 2.

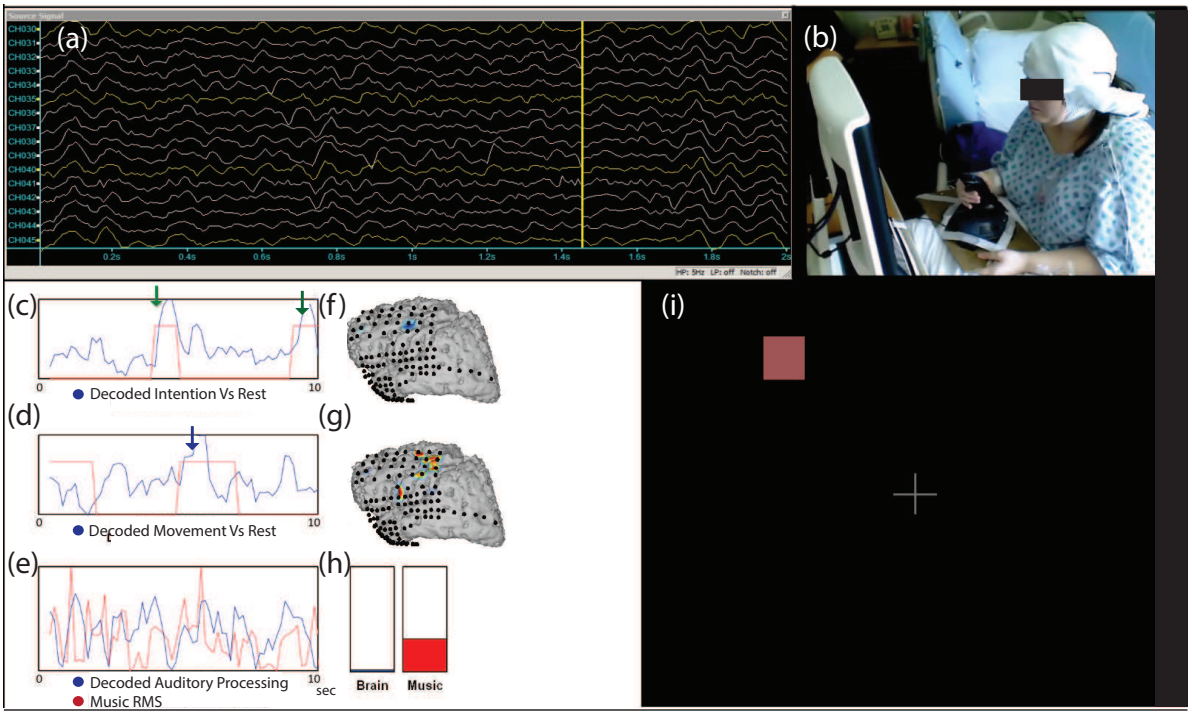

Figure 1. This figure shows the interface for visualizing decoder outputs in real time. Panel (a) shows the real-time viewer for the raw ECoG signals; (b) shows the video feed of the subject performing the task; (c) and (d) show the decoder outputs (in blue) for the intention and movement decoders, respectively, with the respective critical periods marked in red (arrows have been added to emphasize the critical periods' onset times); (e) shows the decoder output (in blue) together with the sound RMS recorded by the microphone (in red); (f) and (g) show the cortical model for this subject and the electrodes that showed significant task-related high-gamma activation in the intention and movement periods respectively; (h) shows the decoded sound RMS and microphone output RMS as bar-gauge visualizations that fluctuate in real time; (i) shows a view of the application screen during the intention period of the center out joystick-task. The decoder traces (c), (d) and (h) scrolled continuously from right to left, with the most recent values plotted at the right-hand edge.

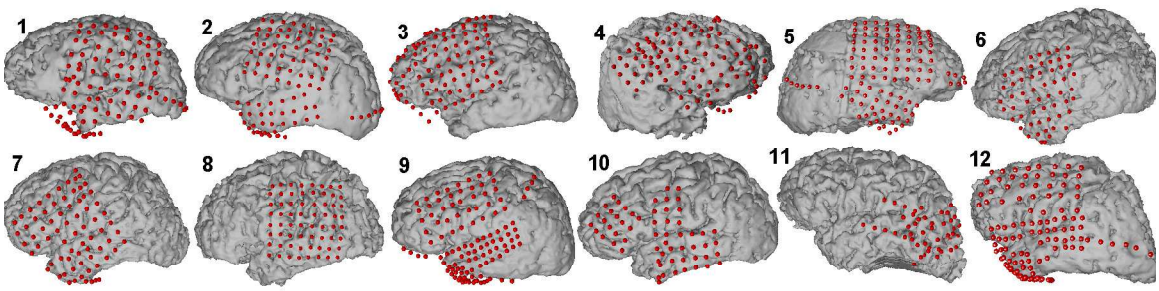

Figure 2. For each of the subjects 1–12, subdural ECoG electrode locations are shown as red dots on a rendering of the subject's three-dimensional surface model of the cortex. The brain model and electrode locations were derived from co-registration of pre-implantation MRI and post-implantation CT.

2.2. Stimulus and Task Design

The cortical functions that we probed in the current study were **intention** (i.e. motor planning in a joystick task), **movement** (i.e. motor execution in a center-out joystick task) and **auditory processing** (i.e. listening to music). This selection was based on recent studies that demonstrated, in retrospective analyses, the relationship with parameters of these functions with brain signals [37, 36, 23, 24, 31, 38, 18, 19, 21, 22]. Recordings also included a **rest** condition.

2.2.1. Center-out joystick task We examined intention and movement with a standard center-out joystick task. The subject was equipped with a joystick and an LCD monitor placed at eye level at a viewing distance of 55–60 cm. The monitor had a built in eye-tracking device that was calibrated to the subject's eyes at the start of every new session.

The task consisted of discrete trials that lasted up to 6 seconds each. These trials were performed in runs of 30 trials, after which the subject could take a short break. In each session, we collected at least 10 runs, for a total of at least 300 trials. A fixation cross was displayed at the center of the screen, at all times, during a run, and the subject was asked to maintain fixation on the cross. The time course of a typical trial is shown in Figure 2.2.2(a). Each trial began with the presentation of a colored square for one second at the target location—this location was chosen randomly and independently on each trial from one of eight positions located equidistant from the center, at 0, 45, 90, 135, 180, 225, 270, and 315 degrees. This one-second period was considered the motor planning or **intention period**, after which a small spherical cursor was presented at the center of the screen. The presentation of the cursor was the cue for the **movement period** to begin: while maintaining central fixation, the subject had to use the joystick to move the spherical cursor until it hit the square target. When the cursor hit the target, it turned green to indicate a hit. A maximum of three seconds was allowed for this movement, after which there was a one-second **rest period**, with only the fixation cross visible. The trial was aborted if, at any time, the subject's gaze shifted away from the fixation cross by more than 20% of the screen height for more than 500 msec.

2.2.2. Listening task The task used for decoding auditory processing involved listening to 3–5 minutes of music—either *The Wall* by Pink Floyd, or *All Right Now* by Fire and Ice. The music was binaurally presented to each subject using standard speakers (50 Hz - 20 kHz audio bandwidth). The sound volume was adjusted to a comfortable level for each subject. Each piece of music was followed by two minutes of rest, during which the subject was asked to relax with eyes open, and avoid movement or speech. The typical timeline is shown in Figure 2.2.2(b). The task did not involve any visual or motor engagement.

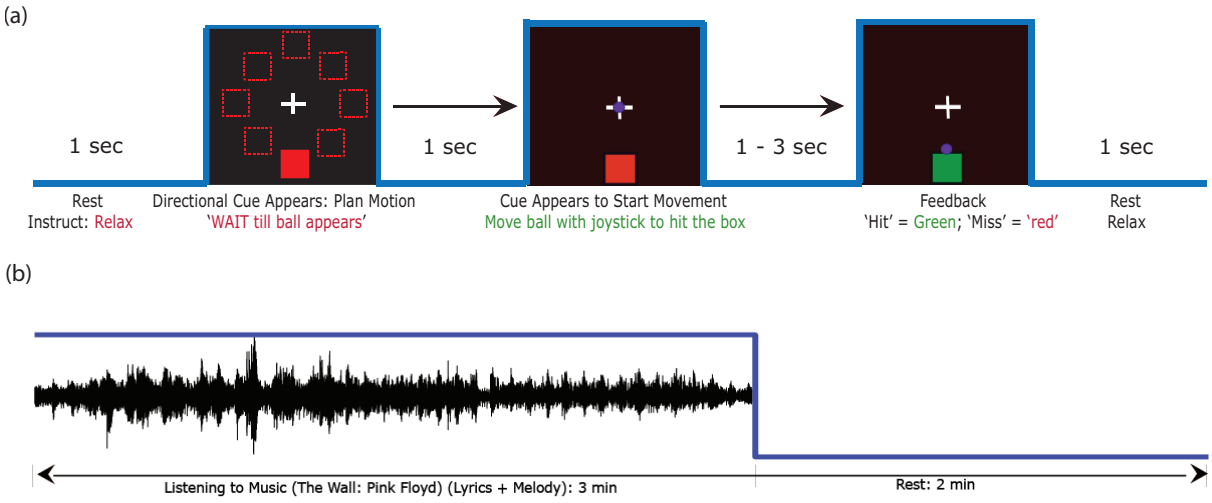

Figure 3. Panel (a) shows the time course of a typical trial in the center-out joystick task that we used to investigate intention-to-move and movement. The subject fixated on the central cross throughout the trial. The 1-second intention period began with a directional visual cue (red square) presented on the screen at the target location. The 1- to 3-second movement period began when a cursor appeared in the center of the screen. This cued the subject to begin joystick movement to guide the cursor to the target. If the target was hit, it turned green. Afterwards, there was a 1-second rest period before the next trial began. The rest period also began if the cursor did not hit the target within a 3-second period. Panel (b) shows the time course of the listening task used to investigate auditory processing: the subject listened to 3–5 minutes of popular music, then rested quietly for 2 min.

2.3. Data acquisition

To implant the ECoG grids, one or two subdural 8 × 8 grids of platinum-iridium electrodes were placed subdurally on the surface of the brain, together with one or more strips that consisted of 4–8 electrodes configured in a single row. The electrodes were of one of two types: those supplied by Ad-Tech (Ad-Tech Medical Instrument Corporation, Racine, WI, USA) had a 2.3 mm exposed recording surface and an inter-electrode distance of 1 cm; those supplied by PMT (PMT Corporation, Chanhassen, MN, USA) had a 3 mm exposed surface and an inter-electrode distance of 6 mm. The grids were placed solely as required for clinical evaluation, without any consideration of the research. They were typically implanted for a period of 4–7 days.

Signals from these ECoG electrodes were fed simultaneously to the research and the clinical systems via splitter boxes (Ad-Tech), with separate ground connections for the two systems. The research system was connected only during research measurements, and collected ECoG signals at 1200 Hz sampling rate using eight synchronized FDA-approved 16-channel g.USBamp amplifier/digitizer units (g.tec, Graz, Austria). Electrodes that were distant from the expected seizure focus, and expected to be inactive or least eloquent with regard to the investigated functions, were selected as the reference and ground electrodes. Data collection, stimulus presentation, and synchronization

of data, stimuli, and joystick/eye movements were accomplished using the BCI2000 software platform [41, 42].

To monitor and enforce fixation, we recorded subjects' eye gaze using a Tobii T60 eye-tracking monitor (Tobii Tech., Stockholm, Sweden) that was positioned at eye level 55–60 cm in front of the subject and was calibrated for each subject at the start of each experimental session.

We recorded the music that was played during the real-time sessions using a dynamic (moving-coil) microphone (frequency response of 50 Hz – 15 kHz). The microphone was electrically isolated and had a unidirectional (cardioid) pickup that minimized unwanted background noise. The recorded sound was fed back into the BCI2000 software pipeline, in real time, parallel to the ECoG data acquisition, to visualize the auditory decoding performance for that piece of music.

2.4. Electrode localization

To localize the electrodes, we collected different types of imaging for each subject: (a) pre-implantation magnetic resonance images (MRI) (T1-weighted coronal SPGR slices, 1 mm width, imaged using a GE 1.5T scanner); (b) intra-operative photographs of the exposed cortical surface before and after grid placement; (c) post-implantation computer tomography (CT) scans (1 mm slice width, skin to skin); and (d) post-implantation lateral and frontal X-Ray images.

We localized the electrode grids by first extracting the three-dimensional anatomical information from the subject-specific pre-implantation MRI images. We then performed a landmark- and volume-based co-registration with the post-implantation CT images using CURRY software (Compumedics, Charlotte, NC, USA). This allowed the projection of grid electrodes onto the modeled cortical surface, including those that were hidden from view during the surgery. For subject 5, post-operative CT images were not available. We localized the grid for this subject with pre-implantation MRI and the post-implantation X-Ray images using the method described in [36]. Electrode locations were verified visually, using the intra-operative grid implantation photographs, based on vascular landmarks. The cortical model and the electrode coordinates were then exported in MATLAB-readable formats for subsequent processing and visualization.

The electrode coordinates were also transformed to the Talairach coordinate system (based on the Anterior Commissure as the origin) and an approximate Brodmann-area assignment that was determined for each electrode using the Talairach Daemon [43, 44]. The resulting approximate locations were used for identifying neuroanatomical cortical structures relevant for the functions under study. This helped to screen for electrodes with (possible) dominant artifactual activations (i.e. electrodes that had significantly high activations during the task, but were in areas far removed from the cognitive process being assessed).

2.5. Functional mapping

For each subject, we established a functional map of receptive language and hand/oral motor function using the SIGFRIED mapping procedure [23]. SIGFRIED compared, at each electrode, the statistical distributions of gamma activation between the resting and activity conditions. The results provided probabilistic maps of significant gamma activation during the functional tasks.

These ECoG- or stimulation-based mapping results allowed us to constrain the choice of electrodes suggested by the subsequent calibration procedures to those locations that were most neuro-anatomically relevant to the task.

2.6. Decoding methods

The common decoding pipeline used in the above phases consisted of preprocessing/feature extraction, followed by either univariate or multivariate decoding.

Preprocessing and feature extraction:

(i) The signals were digitally filtered using an IIR notch filter centered on 60 Hz to remove power-line noise.

(ii) We first excluded channels from the analysis if they were still heavily affected by electrical artifacts such as noise due to lack of electrode contact, non-physiological artifacts such as from monitoring devices or movements, or physiological artifacts such as from paroxysmal activity or seizures.

(iii) The signals were either referenced to the single ECoG electrode that had originally been chosen as a measurement reference, or re-referenced to the common average of all channels (CAR). We evaluated the impact of these two referencing methods during the optimizations of Phase I, and used CAR in subsequent phases.

(iv) The signal was windowed using a rectangular window sliding in steps of 50 msec. For our decoding of intention or movement, one such window was extracted for each trial, centered 500 msec after the start of the relevant period of the trial, for each of the three periods of interest (intention, movement and rest). During optimization, window lengths of 150, 250, 350, 450, 550 and 650 msec were compared; we selected 650 msec for subsequent phases of the study. For our decoding of the temporal envelope of sound RMS, we compared the use of window lengths of 150 msec - 650 msec windows as well, where 650 msec, with a constant lag of 200 msec, was selected as an optimal setting for the study. The lower range of window lengths to be tested was selected based on the commonly used window lengths in similar ECoG studies [13, 23]. Our choice for the higher range of the window length was somewhat more arbitrary, but reflected the trade-off between inclusion of more data for decoding and the requirement for rapid updates and feedback in a real-time system.

(v) The amplitude spectrum was then estimated for each channel and each time-window using an auto-regressive model optimized via Burg's Maximum-Entropy Method (MEM)[45]. We compared model orders of 10, 25, 50 and 100, and used a model order of 50 for all subsequent real-time experiments. Spectral amplitudes were estimated using the resulting model and summated over the range 70–100 Hz. The result was a single estimate of high-gamma activation per channel, per time-window. A growing number of recent studies are suggesting that activity in the gamma band reflects a broadband phenomenon rather than a narrowband oscillation [36, 37]. In our study, we chose the particular frequency band (70-110 Hz), because it avoids the 60 Hz and 120 Hz frequencies that may be affected by line noise. Also, 70 Hz is far beyond the highest frequency (about 30 Hz) at which low-frequency oscillations can be observed; and 110 Hz is much lower than the noise floor of the amplifier system (about 200-250 Hz).

(vi) The high-gamma amplitudes were then represented on either a linear or a log scale. After initial evaluations, we chose the log scale for representing the ECoG features in subsequent analyses.

Multivariate Decoding: The decoding component mapped extracted features to estimates of engagement in the relevant neurocognitive function as compared to resting state. The final output signal was computed by combining the high-gamma activation features across channels using a weighted linear summation.

For the music listening task, multiple linear regression was used to regress the root-mean-square (RMS) of the music RMS (computed from the music file in sliding windows of the same length as the window used for ECoG high-gamma amplitude estimates) against the ECoG high-gamma estimates.

For the joystick task, detectors for the intention state and the movement state were built by considering the two corresponding binary classification problems, intention vs. rest and movement vs. rest. For completeness, we also show results for the intention vs. movement problem. Linear weights for these binary classification problems were obtained with the stepwise (SWLDA) multilinear regression method, which determined the weights of the linear function so as to minimize the squared error between the output estimates and labels of the classes (-1 and +1).

Univariate decoding: Previous studies [37, 35] showed that some functions can be tightly spatially localized on the ECoG grid. We found that our functional mapping results from the SIGFRIED system [23] often corroborated this observation. On the assumption that the single most functionally relevant spatial location is unlikely to change as a function of time at the spatial scale at which we record, it seemed plausible that the selection of a single electrode might provide good performance while generalizing well across sessions and contexts, and minimizing the risk of overlap between decoders intended to reflect different brain functions.

To test this hypothesis, we compared the multivariate decoder against a univariate decoder, in both our optimization (Phases I–II) and where possible in our real-time test (Phases III–IV). The univariate decoder was based on manual selection of a single electrode location, as selected by the researcher who integrated the following multiple sources of information:

(i) topographical maps of the coefficient of determination (r^2) that were calculated, for each electrode, between band power of the ECoG signal and the state of task-engagement (see Section 2.7);

(ii) anatomical structures covered by the electrodes, as obtained from various grid localization methods, such as intra-operative photographs (Figure 2.6(a,b)), pre-operative and post-operative CT and MRI imaging, X-ray and MRI imaging (Figure 2.6(c,d)), and surgical notes (Figure 2.6(f)) (described in Section 2.4);

(iii) topographical maps of Brodmann areas that were derived by rendering the patient-specific 3D cortical anatomy with electrodes in the Talairach coordinate space (Figure 2.6(e)) (described in Section 2.4);

(iv) functional activity maps obtained from SIGFRIED (Figure 2.6(g,h)) or electrical cortical stimulation (see Section 2.5);

(v) areas masked/activated by epileptiform activity as defined by clinicians.

2.7. Decoding assessment

We assessed decoding performance in the following ways:

- Assessment of relevance of individual features: to assess the extent to which ECoG activity at one or more locations contributed to solving a prediction problem, we used the coefficient of determination (r^2). This metric reflects the fraction of the variance in the predictor's output that is explained by the differences in the target labels. For binary classification problems, the labels were arbitrary (-1 and +1), whereas for our listening task, the labels were the RMS values of the music envelope, computed from the music file using sliding windows of the same length as the windows used for feature extraction from the ECoG. When we assessed the significance of an r^2 value, we used Bonferroni-correction to account for multiple comparisons.
- Assessment of predictor performance in the music listening task: this was quantified by computing the Spearman correlation coefficient ρ between the predictor output and the running RMS values of the music RMS.
- Assessment of trial-based classifier performance: one measure of intention and movement prediction performance is classification accuracy (CA) of a binary classifier. To compute CA, we extracted just one discrete exemplar of each class (intention, movement and rest) from each trial of the center-out joystick task. CA is given by the proportion of these exemplars that the classifier assigns to the

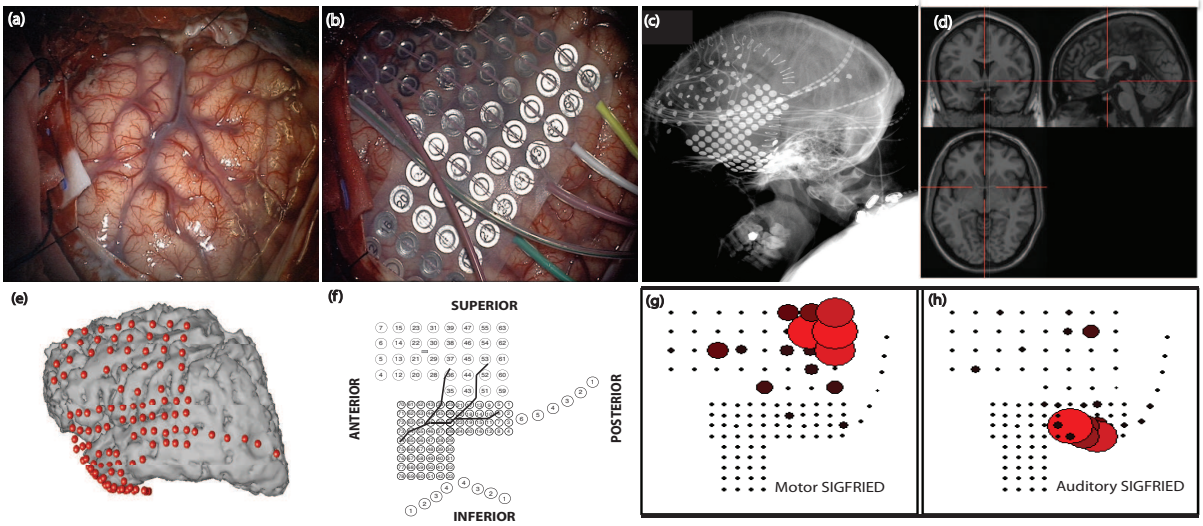

Figure 4. This figure summarizes some of the ancillary sources of information used by the authors to make an informed choice of functionally relevant electrodes when employing the univariate decoding approach. Panel (a) shows an intra-operative photograph taken before grid placement; panel (b) shows an intra-operative photograph with grid placement; panel (c) shows grid positions from a post-implantation CT; panel (d) shows anatomical information from a pre-operative MRI; panel (e) shows a 3D cortical model rendered from a pre-operative MRI and a post-implantation CT, from which approximate Brodmann-area assignments are made by the Talairach daemon; panel (f) shows surgical sketches and notes; panel (g) shows the SIGFRIED mapping for a complex hand motor task, and (h) shows the SIGFRIED mapping for a listening task.

correct class in a binary subproblem (intention vs. rest, movement vs. rest, or intention vs. movement). It is estimated by training the classifier on one subset of the exemplars and testing it on another non-overlapping subset—either repeatedly by cross-validation, or via a single training/test fold, as described in section 2.6.

- Assessment of intention and movement predictors in real time: in a realistic brain-monitoring application, classifiers for each brain state must be trained without complete prior knowledge of the classes against which their output must distinguish itself. For example, if we want to build a monitoring system that includes a gauge labelled "intention to move the hand," then this gauge should ideally distinguish intention to move the hand from all other possible states (i.e. from rest, and from actual movement of the hand, but also from the intention to move other body parts and from the performance of an arbitrary range of other mental tasks). We also assume that the outputs of a monitoring system would not be assessed at single discrete pre-defined time-points (as assumed in an assessment based on CA) but as continuous traces (or at least, traces that are updated in small (e.g. 50 msec)

increments).

Therefore, we wished to ask: given a classifier trained to distinguish a certain brain state (say, intention) from the resting state, how well does it distinguish *all* the time windows of the test set that could be considered to reflect that state (intention), from *all* other time windows (non-intention)? To quantify this, we labelled each time-window as *intention* if the majority of the window's samples occurred during the intention period, and as *movement* if the majority of the window's samples occurred during a movement period. We judged the detector for a brain state to have predicted that state when its output exceeded the 95th percentile of its previous activity, computed over a 10-second sliding window (the same length of time that our visualization window displayed at any one time). We then computed sensitivity and specificity of these detectors for intention vs. non-intention windows, and for movement vs. non-movement windows:

$$\text{sensitivity} = \text{true positives} \,/\, (\text{true positives} + \text{false negatives})$$

$$\text{specificity} = \text{true negatives} \,/\, (\text{true negatives} + \text{false positives})$$

In the optimizations of Phase I, where we sought to find the optimal feature extraction parameters, the split into training and test subsets was performed repeatedly in a 10-fold cross-validation procedure: the data were divided into 10 parts, each tenth took a turn at playing the role of the test set for a classifier trained on the remaining 9 tenths, and the resulting CAs on these 10 test sets were averaged. We ensured that the trials were temporally contiguous within each test fold—that is, we performed *blockwise* cross-validation—to avoid potential inflation of the CA estimates due to non-stationary effects in the data [46].

Cross-validation in 10 independent folds was only possible for automatic classifiers, i.e. for our automated multivariate procedures. Our univariate classifiers involved human judgment, which could not be guaranteed to be independent on 10 folds. Therefore, for comparison of univariate and multivariate methods during Phase I, we used an *offline single-fold* training and test procedure for both predictor types, whereby the first 80% of the data (counting chronologically) were used for training and the subsequent 20% were used for testing. This schedule was used both for classifier performance and for assessment of music listening prediction. During phase III, the single-fold method was also used to provide a preview of decoder performance based on the calibration data.

In assessing generalization from phase I to phase II, and from phase III to phase IV, one whole session was used for training and the subsequent complete session was used for testing.

3. Results

The results of this study are presented below. They address the following questions:

(i) What is the best combination of parameters settings for decoding intention and movement states in the joystick task, and how sensitive is the decoder to variation

from this setting? Similarly, what is the best combination of parameter settings for decoding auditory processing?

(ii) How well can we expect the decoders to generalize to a second session? Furthermore, will decoders trained on a trial-based system, with brain state classes considered separately against the rest state, perform well in a continuous real-time setting where they must distinguish their preferred state from both rest and from other non-preferred non-rest states?

(iii) Which decoding approach, the automated multivariate method or the expert-guided univariate method, provides better decoding performance?

(iv) How many training trials are required to achieve good multivariate decoding performance?

(v) How well will our decoders generalize to a second real-time session in which the tasks are performed simultaneously? Will they perform well according to a real-time (continuous rather than trial-based (see Section 2.7)) criterion?

3.1. Parameter optimization and sensitivity

Our first question concerned the optimal parameter setting for intention and movement decoding, and the sensitivity of the decoder to parameter variation. As described in Section 2.6, we evaluated the impact of several types of processing: referencing strategy (CAR/no CAR), window length (150, 250, 350, 450, 550 or 650 msec), AR model order (10, 25, 50 or 100), and gamma transformation (log/linear). We addressed these questions using the automated multivariate decoding approach and repeating the 10-fold cross-validation procedure for every possible combination of parameters. The classification accuracies were averaged across the first six subjects. They indicated the following optimal settings:

Subproblem	CAR	Log Band Power	AR Model Order	Window Length (msec)
Intention vs. Rest	Yes	No	50	650
Movement vs. Rest	Yes	Yes	10	650
Intention vs. Movement	Yes	Yes	10	650

Figure 5 shows the sensitivity of the decoders to the variation in these settings. The optimization results indicated that: first, the use of a log bandpower scale improved classification accuracy in movement vs. rest and movement vs. intention, but not in intention vs. rest, albeit all of these effects were very small. Second, the use of CAR spatial filter improved the classification accuracies in all subproblems by about 5 percentage points. Third, a lower AR model order of 10 was preferable for movement vs. rest and movement vs. intention, while a higher model order of 25 or 50 was preferable for intention vs. rest. We chose to boost the performance of the lower performer of our two detectors by selecting the model order of 50. Fourth, longer window lengths improved classification accuracy in all subproblems, best performance being obtained with a length of 650 msec. Based on these results, we settled on parameter setting for all subproblems included the use of log band power, from CAR-rereferenced data, with an AR model order of 50 and a window length of 650 msec.

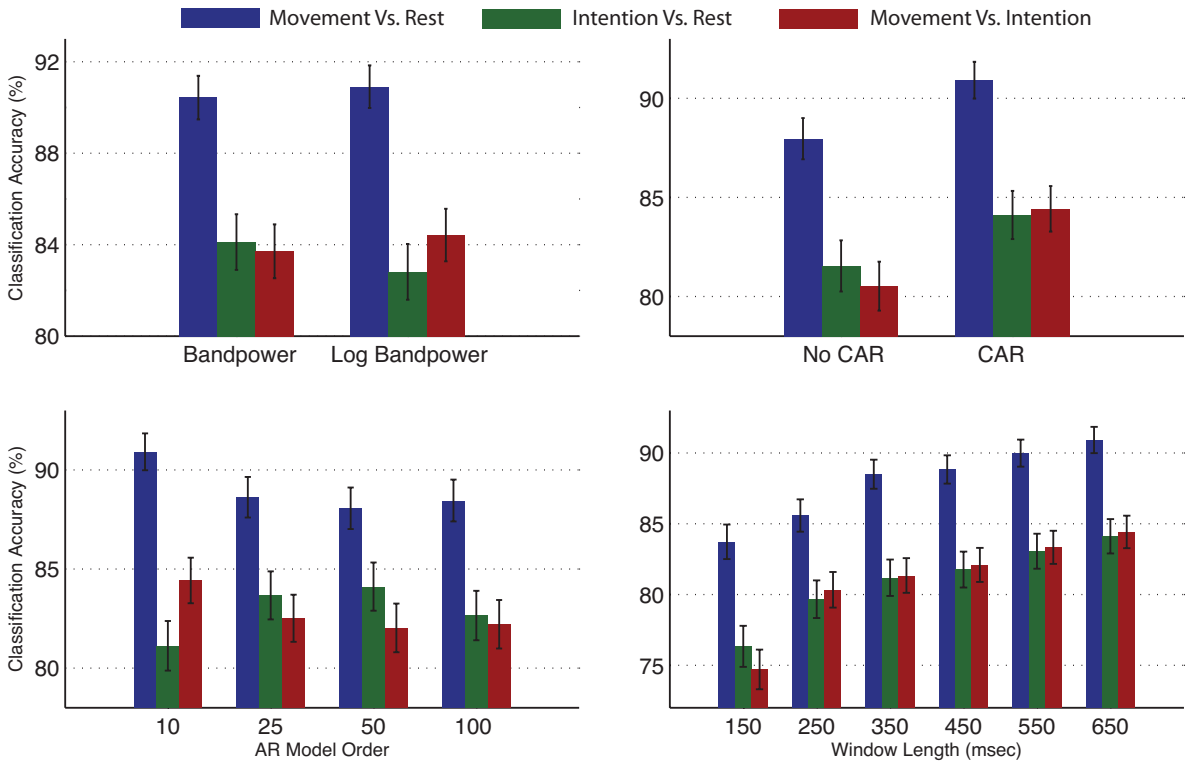

Figure 5. This figure shows cross-validated classification accuracies (%) averaged across subjects 1–6 with error bars indicating ±1 standard error, for solving three binary classification subproblems with a multivariate classifier. Chance performance would be 50%. From left to right within each group of bars, the subproblems are movement vs. rest (blue), intention vs. rest (green), and movement vs. intention (red). To illustrate the sensitivity of the decoder to the preprocessing parameters, each of the panels shows the effect of varying one of the parameters away from the global optimal combination.

We also addressed similar issues of parameter optimality and sensitivity for the listening task. For technical reasons, the same set of pre-processing parameters had to be used for all decoders simultaneously. We were satisfied that our CAR, log and model-order settings chosen in the joystick-task optimization represented a good general way of extracting high-gamma features. Therefore, this question boiled down to verifying whether our chosen window-length, in combination with these parameters, was also optimal for the listening decoder. An additional parameter for the listening task is the temporal lag (between sound delivery and brain response) at which the listening decoder should be trained and the results assessed. To address this, we fixed the CAR, log and model order settings at the optimal values determined for the joystick task, and evaluated the six different window lengths in combination with temporal offsets from -2 to +2 sec in steps of 20 msec. The correlation between neural response and music stimuli at varying lags and window lengths is shown in Figure 6(a) for a univariate decoder. The correlation was observed to be highest (rho = 0.34) for the longest window length of 650 msec and peaked at a lag of approx. 200 msec when averaged across subjects

(50 msec of which can be accounted for by the software's latency in producing stimuli).

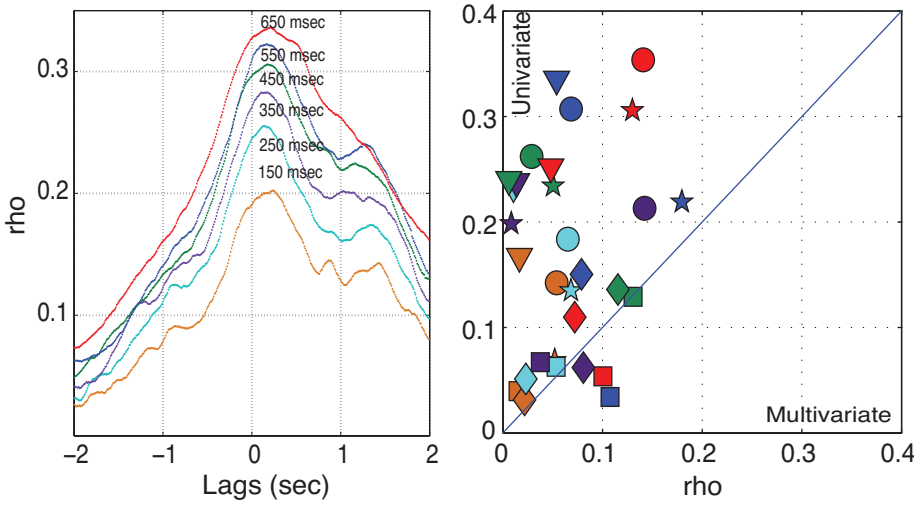

Figure 6. (a) Correlation coefficients of univariate auditory features and music RMS, across subjects 1–6, at leads and lags of 0 to 2 seconds and window lengths 150–650 msec (b) Correlation coefficients for subjects 1–6 (each represented by a different symbol shape) and all window lengths, at a constant lag of 200 msec, estimated by a multivariate and a univariate regression. Color represents the window length used for the analysis, corresponding to the colors in (a)

3.2. Session-to-session generalization

We then wished to assess the ability of the decoders to generalize to a second day. This is an important issue, because changes in arousal or other brain states on a timescale of hours or days may be substantial. Classifiers trained on the first day's session were tested on second-day data, which were available (still from tasks performed separately) for 4 of the original 6 subjects. This performance was compared against the performance from a single-fold offline analysis of the day-1 data. The results for intention and movement decoding are shown in Figure 8 (b), where each symbol represents a different combination of subject and subproblem. Symbol shape denotes the subproblem as indicated in the legend. There is a drop in performance of 6.8 percentage points (on average across all subjects and subproblems) when transferring a trained multivariate classifier from one day to the next. For the univariate decoder, this average drop is larger (11.4 percentage points). The paired t-test between the multivariate decoding accuracies from day 1 and day 2 result in a p-value > 0.05, both for motor intention and motor execution. This shows the robustness of a multivariate decoder for day-to-day transfer. However, the paired t-test between the univariate decoding accuracies from day 1 and day 2 result in a p-value < 0.05 for motor intention and p-value > 0.05 for motor execution. This shows that atleast for the motor intention processes, the univariate decoder is less robust for day-to-day generalization.

We also wished to confirm our hypothesis that individual decoders that were trained on separate tasks would be able to deliver reliable outputs, separately as well as simultaneously. The continuous decoding performance was quantified in terms of decoder sensitivity and specificity for both multivariate and univariate analysis. Sensitivity and specificity measurement on the continuous decoding quantifies the ability of the decoder to detect or reject the intention (or movement) states not only in contrast to the marked rest states, but also in contrast to the other ongoing states (inter-trial, movement (or intention) and feedback) in an analysis where one data point was a single time-window rather than a whole trial. These have been presented in Figure 7 (left panel). Across subjects, the multivariate and univariate decoding were found to be equally sensitive for intention and movement decoding. However, multivariate decoding was found to show relatively higher specificity for intention decoding, while univariate decoding showed higher specificity for movement decoding. The r^2 maps for all the subproblems are also shown in the right panel of Figure 7. The running correlation coefficient between decoder output and music RMS was calculated using a 20-second sliding buffer to reflect the variation in auditory system decoding accuracy as a function of time during the song. It is shown in the panel (e) of Figure 7.

3.3. Multivariate vs. univariate decoding

Our next question required a comparison of the automated multivariate and expert-guided univariate decoding methods described in Section 2.6. The comparison was performed using a single-fold analysis of the data from subjects 1–6.

For the listening task, the results are shown in Figure 6(b). The colors denote window size in the same way as indicated in the adjacent panel (a). Each symbol shape denotes a different subject. We see a clear and consistent advantage for the univariate method over our chosen multivariate decoding approach—hence we show results of the univariate approach in panel (a), and use the univariate approach throughout the rest of the study for the purpose of auditory decoding.

The results for intention and movement decoding are shown in Figure 8 (a and b). Across all 6 subjects, all 3 subproblems, and both days, our multivariate decoding approach classified the data with an average of 81.5% accuracy, whereas our univariate approach was 75.8% accurate. Note that the univariate approach could also achieve very high performance levels, depending on the subject—but the variability across subjects was larger (see Figure 8(a)). The multivariate classifier also produced more *consistent* results when generalizing from one day to the other, as we saw above in Section 3.2.

3.4. Amount of required training data

Our next question centered on the impact of the amount of calibration data on decoding performance. The performance of a predictor can generally be expected to improve as the amount of training data increases, although these improvements should diminish

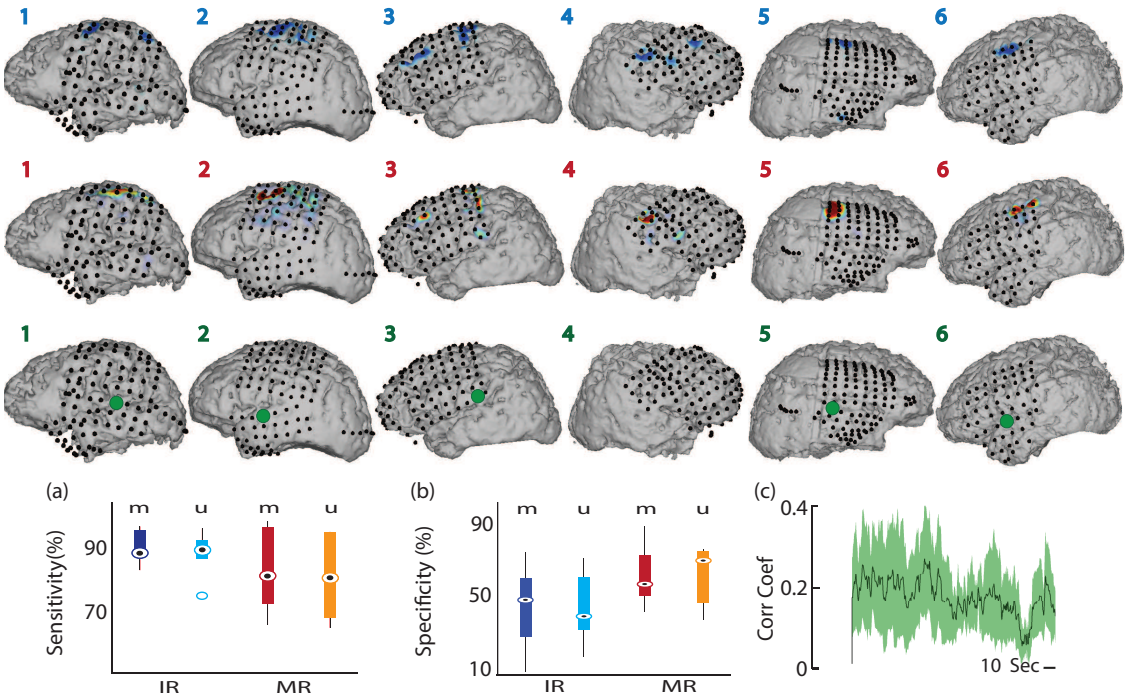

Figure 7. The r^2 maps for intention, movement and auditory processing are shown for subjects 1–6. Top row shows the topographies for Intention, second row shows topographies for movement and third row shows topographies for auditory processing. The electrode selected for univariate auditory analysis is marked with a green circle. Lower panel (a) and (b) show the sensitivity and specificity for the simulated real-time decoding across subjects 1–6, for both multivariate (m) and univariate (u), intention vs. rest (IR) and movement vs. rest (MR). Lower panel (c) shows the running correlation coefficient for auditory vs. music RMS, across 1–6 subjects, using a 20 sec window at a time.

with larger data sizes. We investigated this issue by computing the trial classification accuracies (CA, as defined in Section 2.7 above) in single-fold assessments where the size of the training fold was either 20, 80, 140, 200, 260 or 320 trials. The results, averaged across subjects for each subproblem, are shown in Figure 9. They display the expected increasing trend, which seems to level off at around 140 trials—slightly less than half the data we gathered.

3.5. Real-time simultaneous multimodal decoding

In subjects 7–12, we examined the real-time performance of the system. Here we quantified how well our decoders that were trained on data from separate tasks (motor intention, motor execution, auditory) generalize to a second real-time session in which the tasks are performed simultaneously. Furthermore, we wanted to determine whether decoders that were trained on data from discrete trials, with brain state classes considered separately against the rest state, performed well in a continuous real-time

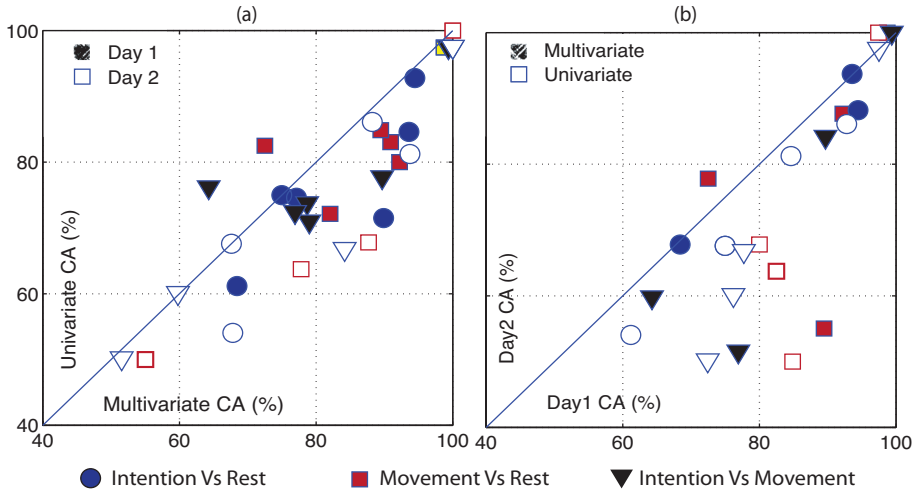

Figure 8. (a) Multivariate vs. univariate decoding classification accuracy for day 1 and day 2 test data of subjects 1–6. Each symbol represents a subject and each symbol shape represents a decoding subproblem as described in the legend. The filled shapes represent results of day 1; the unfilled shapes represent day 2. (b) The multivariate and univariate classification is shown for day 1 vs. day 2. Filled shapes represent multivariate results and the unfilled shapes represent the univariate results. (c) Multivariate vs. univariate decoding performance shown by correlation coefficients of auditory features and music RMS for subjects 1–6 (excluding subject 4).

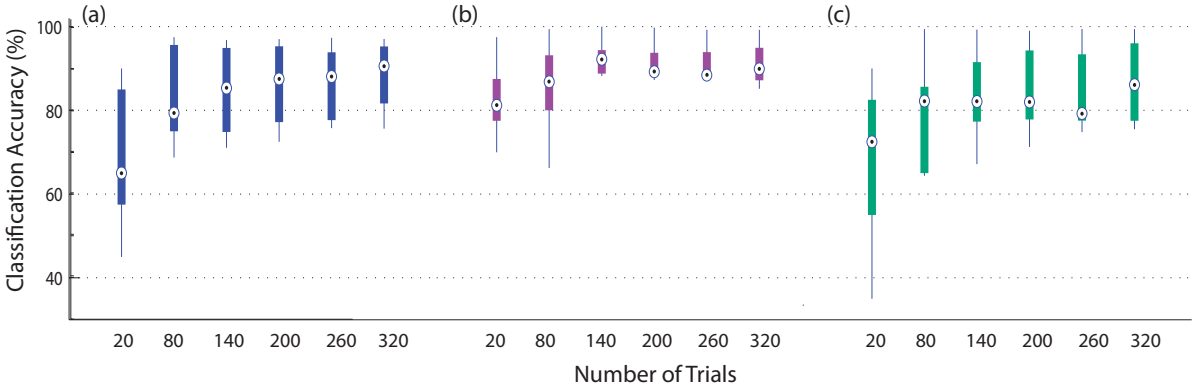

Figure 9. The effect of the amount of the training data on decoding performance is shown for the 3 motor-based decoding subproblems: (a) Intention vs. Rest (b) Movement vs. Rest (c) Movement vs. Intention.

setting where they must distinguish their preferred state from both rest and from other (non-preferred/non-rest) states?

Based on the optimization outcomes from Phase I and II, we used the multivariate decoder for intention and movement decoding and a univariate decoder for auditory decoding. Where possible, we also performed a second real-time run with the univariate decoder. This was possible for subjects 9, 10 and 12. Subject 11 lacked coverage of

motor and pre-motor areas. For the remaining two subjects (7 and 8) a second run was not possible due to time constraints, so we re-performed the analysis retrospectively—however, the order of transfer from training to test was always strictly causal and no re-optimization of the decoder was allowed.

An example of the real-time decoding performance can be seen in Supplementary Movie (S1). The movie was captured (with the patient's consent) using the Camtasia software application, which captured live video footage of the patient performing the task and also captured the computer screen the contained real-time visualizations for the patient and for the investigator.

The task-related r^2s were estimated for all the electrode channels of the training data. The results are presented in Figure 10. The overall classification accuracy for these subjects is shown in panel (a). Again, we found that multivariate decoding provided higher performance for motor decoding (median classification accuracy across subjects was 80.4% for intention decoding and 86.8% for movement decoding). Panel (b) shows univariate auditory decoding performance for these subjects using the non-parametric Spearman's correlation that was calculated between gamma features and music RMS. These results demonstrate that it is possible to accurately decode movement intention, movement, and auditory processing concurrently and in real time.

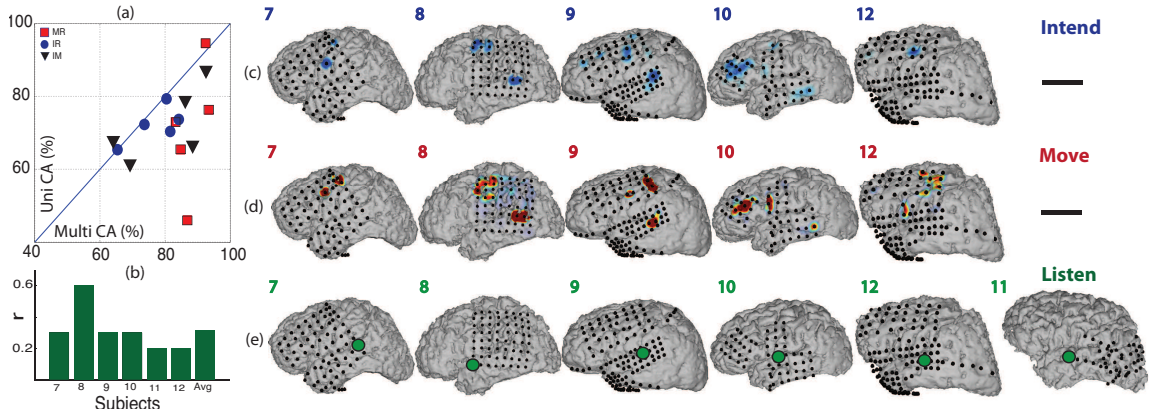

Figure 10. (a) Multivariate and univariate classification accuracy (CA) is shown for motor decoding for subjects 7–12 as tested on data from day 1. Each symbol represents a decoding subproblem. (b) The correlation coefficient (r) for auditory decoding vs. music RMS from day 1 is shown for subjects 7–12. Across all subject, the average r value was 0.3. (c, d, e) are the r^2 maps for intention, movement and auditory tasks, for subjects 7–12. Subject 11 did not have motor coverage.

Figure 11 presents a quantitative analysis of the trial-based classification performance for subjects 7–12 during calibration (Phase III) and real-time testing (Phase IV). The figure is comparable to figure 8, which showed results in the same format for subjects 1–6 (Phases I and II). The difference between these analyses is that only subjects 7–12 performed the two tasks simultaneously in their second session. As in Section 3.3, we find that multivariate decoding (average accuracy 76.5%) resulted in higher performance than univariate decoding (70.4%) of intention and execution of

movement.

In common with our Phase I–II results, we observe a drop in performance when we force our decoders to generalize from one day to the next. However, this time the drop in performance is larger for the multi-variate decoder (13.9 percentage points) than for the univariate decoder (5.4 percentage points). This is the reverse of the pattern we found before, in Phases I–II. We can interpret this by considering that we are now demanding that our decoder generalize between data sets that are more heterogeneous than before. Before, the decoder was both trained and tested on distributions of high-gamma features measured from the joystick task performed in isolation. Now, the decoders are still being trained on data measured during the joystick task alone, but tested on data measured with a concurrent listening task. Training and test set distributions are therefore different. As we pointed out in Section 2.6, our univariate decoder uses only one electrode, and hence only captures information from a highly localized cortical region. It will therefore suffer little interference from the activity of brain areas that are functionally unrelated to its primary task. By contrast, a multivariate decoder is free to use information from the whole grid, even from brain areas that are functionally unrelated to its primary task, and it may do so if this serves to suppress noise. For example, under constant quiet conditions, a multivariate decoder trained to identify motor intention might incorporate information from auditory cortex electrodes: it may be that auditory-cortex activity provides a good baseline against which to measure fluctuations in premotor cortex activity. This may generalize well, but *only* as long as the conditions remain quiet.

An example of the real-time decoding for subject 12 is shown in Figure 12. It presents the motor and auditory processing real-time decoding as estimated on separate tasks as well as during simultaneous tasks, predicted by the univariate approach. Panel (a) and (b) show the univariate decoding on decoders that were trained separately and tested real-time separately on each of the intention-movement and listening tasks. Panel (c) shows the univariate decoding obtained when the intention-movement was performed simultaneously with the listening task, on the same day. The time segments during which *intention to move* was being expected, has been indicated by semi-transparent blue back-panels in (a) and (c). While the time time segments during which *movement execution* was being expected have been indicated by semi-transparent red back-panels in (a) and (c).

Finally, sensitivity and specificity was performed with the Phase-IV simultaneous decoding data. The univariate analysis was performed in parallel as a backup alternative. The overall classification and the sensitivity and specificity outcomes are shown in Figure 13 for subjects 7–12. This result shows the ability of the decoders to not only detect a change in state aptly but also the ability to reject rest and other changes from that state. *This was the first indication that the decoding features apparently are quite unique for detection of intention, movement.* For example, even though the training data for intention vs. rest did not include movement periods, the real-time decoding based on

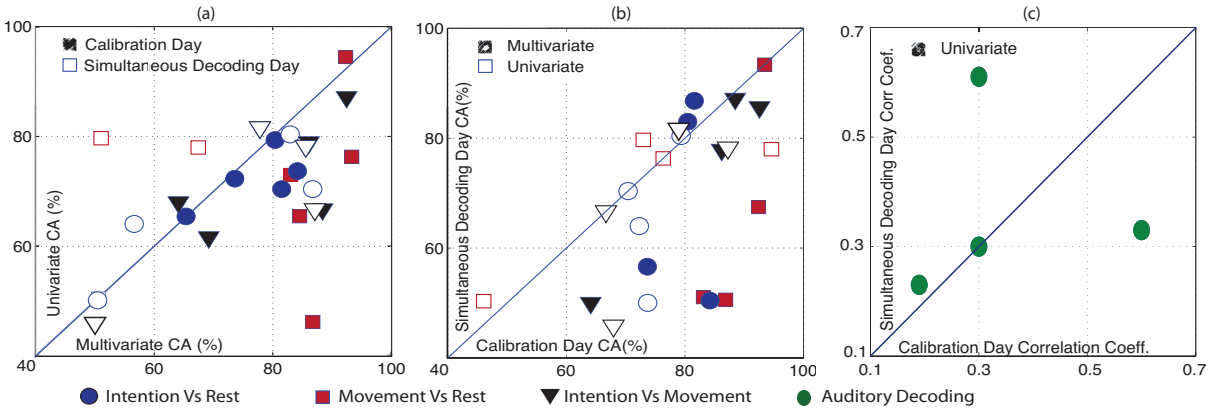

Figure 11. Simultaneous decoding performance and generalization results for subjects 7–12: (a) Multivariate vs. univariate intention of movement and movement execution decoding classification accuracy for data from the calibration day and for data from the subsequent day on which simultaneous decoding took place. Each symbol shape represents a different subproblem as indicated in the legend. The filled shapes represent results of day 1; the unfilled shapes represent day 2. For each day and each subproblem, multiple symbols represent multiple subjects. (b) The same results as in (a) are displayed here in a different manner to highlight the difference in performance between the calibration day vs. the simultaneous decoding day. Filled shapes represent multivariate results and the unfilled shapes represent the univariate results. (c) The auditory decoding performance, measured by correlation coefficient of the decoded signal vs. the music RMS, is shown for calibration day when the listening task was performed separately, compared to simultaneous day when the listening task was run simultaneously with the intention-movement task. Note: subject 11 did not have motor coverage and subject 10 did not perform the 2 tasks simultaneously in real-time in any one session.

the intention vs. rest decoder was able to contrast intention and movement sufficiently well.

4. Discussion

In this study, we set out to develop, implement and test algorithms and procedures that allow us to decode from concurrently performed tasks, within the structure of a common hardware and software framework. Previous studies [47, 22, 37, 12, 23, 16, 38, 35] have demonstrated retrospectively that individual brain processes, such as movement, movement planning or auditory processing can be reliably decoded from ECoG signals. However, it was not clear to what extent these decoding models could be applied in real time or generalize over time. It was also unclear whether they are independent enough to be transferred to more unpredictable multi-modal decoding environments. The use of high-gamma power modulation for such decoding has been found especially useful and so we concentrated only on this feature in the current study. The modulation in high-gamma power correlates with the firing activity of large neuronal populations. At a smaller scale, the firing of individual neurons has been found to be a good predictor

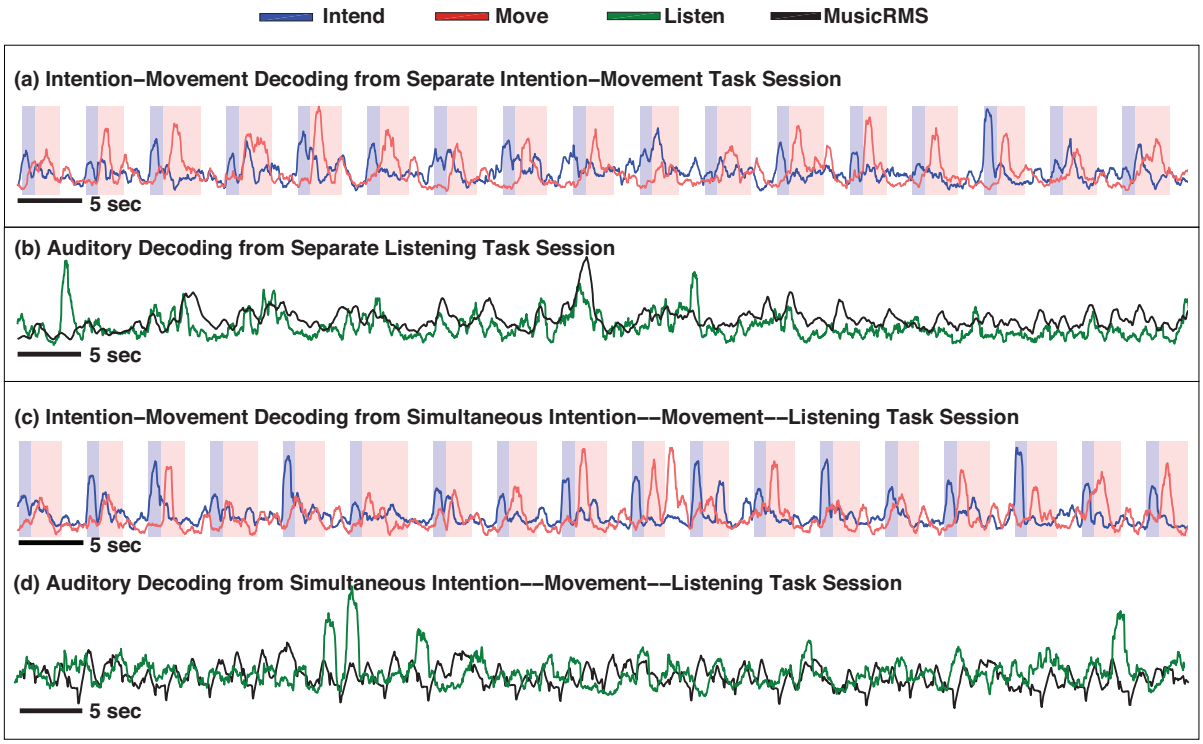

Figure 12. This figure shows an example of real-time separate as well as real-time simultaneous decoding for subject 12 as captured with the univariate approach. The decoded neural output and the expected output are shown for intention-movement task in (a) and (c). The decoded neural output for listening task is shown in (b) and (d). The semi-transparent blue back-panels in (a) and (c) denote the segments during which *intention to move* was being expected and the red back-panels in (a) and (c) denote time segments when *movement execution* was being expected. The decoding results in (a) was obtained when the intention-movement task was executed separately. Decoding in (b) was obtained when the listening task was run separately. Sub-figure (c) shows the real-time intention and movement decoding when the task of intention-movement and listening to must were run simultaneously. Sub-figure (d) shows the auditory decoding obtained during that time.

of movement execution, movement planning and movement kinematics [29, 30, 48, 49]. So far, however, it has been difficult to use such systems long-term without frequent re-calibration [31, 49]. ECoG high-gamma features, on the other hand, strike a good balance between invasiveness, decoding quality and stability [4, 35]. As recording technology continues to miniaturize and implantation techniques continue to improve, the barriers to long-term implantation of ECoG grids can be expected to decrease. In the last few years we have seen many advances in the development of ECoG electrodes towards a smaller-scale, higher-density, more-flexible, more-durable, wireless ideal [50, 51, 52, 53, 54, 55, 56, 57, 58]. With continuing improvements, it is plausible that long-term implantation of ECoG systems may become a realistic option for monitoring, functional restoration or possibly even treatment of a wide variety of brain disorders.

Our results demonstrate that the ECoG gamma features allow single-trial real-time

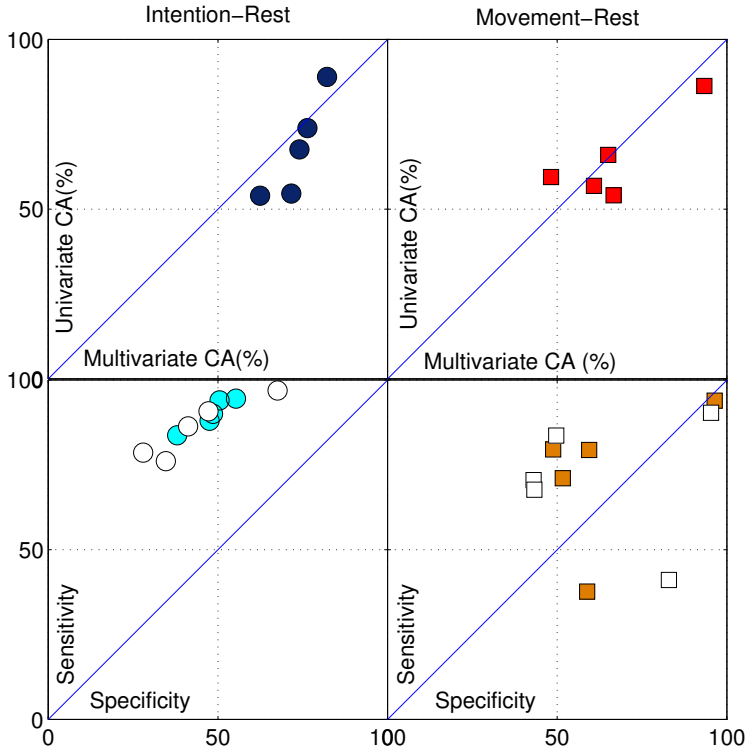

Figure 13. Online continuous decoding results: Upper panels: Multivariate vs. univariate classification accuracy (CA) for intention and movement decoding. Each symbol represents a subject. Lower panels: sensitivity vs. specificity is shown for multivariate as well as univariate approaches.

decoding of movement planning, movement execution and auditory processing with a high sensitivity and specificity. The decoders for each of these states are separable, in the sense that they can be optimized as separate models, yet still perform well when transferred to a multi-modal problem setting: during optimization, each class is distinguished only against rest, and auditory data and motor planning/execution data are gathered in separate sessions; by contrast, in testing, motor planning (for example) is compared against motor execution *or* against rest, and *while* the auditory system is concurrently active. The results are encouraging as they demonstrate the possibility of real-time simultaneous monitoring of multiple brain systems based on a modular combination of simpler single-task training protocols.

The quantitative results of our optimization procedures provide insight into the sensitivity of such systems to certain pre-processing parameters. In particular, the use of a common-average reference improved results substantially across practically all of our analyses relative to data that were referenced to only one reference electrode. Also, temporal smoothing played a key role in improving the results: both our subjective judgments of traces of the kind plotted in figure 12 and the quantitative results shown in figures 5 and 6 indicate that longer temporal windows improve decoding performance. Naturally there is a trade-off between performance and responsiveness in such systems: the longer the window-length, the more accurate the results may be but the less rapidly

the real-time system can respond to temporal dynamics of the brain's state. In addition, it can be seen that the increase in performance is smaller for window lengths longer than 350 msec as compared to the larger increase in performance when we increased the window length from 150 msec to 350 msec. Thus, the optimal setting should rely in part by suggestions from statistical analysis but also be the demands of a particular application. Our results also suggest that the amount of training data could have been reduced to around half (to 150 trials, or 15 minutes not counting interruptions) without an appreciable impact on the quality of intention or movement decoding.

We would like to point out that in its current form this study does not tell us how general our results are across various brain processes or data acquisition setups. Even though we observed similarity in the relationship between decoding performance and parameter selection across our 3 tasks, it is possible that these were suitable only for the particular covert and overt tasks under study (motor intention, motor execution and auditory processing). These may benefit from fine-tuning to the other brain processes being studied and the signal-to-noise ratio of data acquired under different settings.

In Results Sections 3.3 and 3.5, we observed that the multivariate approach generally performs better than the univariate approach in decoding intention and movement. The multivariate decoder generalized better than the univariate decoder to a second day when the task remained the same. However, it suffered a *larger* drop in performance than the univariate approach when generalizing to a later session involving a concurrent listening task. We explained this in terms of the multivariate classifier's ability to use signals from brain areas that are functionally unrelated to its primary task (for example, a decoder for intention-to-move might include information from auditory cortex). This strategy may be effective when the unrelated brain area (the auditory cortex in this example) displays consistent patterns of activity, and thereby provides a useful baseline against which to measure the activity of more directly relevant brain areas (premotor cortex). However, the strategy may become counterproductive as soon as the unrelated brain area becomes simultaneously involved in an unforeseen task (for example, listening to music). If the multivariate decoder performs better than the univariate decoder to begin with, but loses more during generalization to more complex measurement conditions, can we make a final decision as to which is preferable for a multi-system brain monitoring application? One approach would be to assess the decoders' respective performance levels on just the simultaneous decoding session: if we do this we find that their performance is very similar (70% accuracy for multivariate, 68.5% for univariate, averaged across subjects and subproblems). However, we must remember that these results were obtained with only *one* known form of competing information processing (listening to music) beside the primary joystick task. If we were building a brain monitoring system, we would want it to perform well even when the brain had to process yet another kind of information (for example, monitoring the subject's intention to move while background music is playing *and* the subject is experiencing pain). In the translation of our decoders from a controlled experimental study to a more realistic brain-monitoring application, it seems likely that a univariate

approach might be more suitable in the context of a larger number of unforeseen interfering brain processes.

It is worth noting that the multivariate method may be more attractive in practical situations: multivariate prediction is a semi-automated method, while the univariate method as employed here is an approach that at least in part depends on subjective judgments. As the process depends on a very large number of variables (e.g., location of the grid, location of the epileptic focus, brain system modality, results from other functional mapping, or other characteristics of the subject, see Section 2.6), it has been difficult, and will likely remain difficult, to formalize the process by which the different sources were used in the decision process.

Even when we want to decode rich information sources distributed over complex brain networks, a highly localized approach such as our univariate decoding strategy may still yield good results: Wessberg et al. [34] observed in non-human primates that motor control signals from single-neuron electrodes appeared concurrently in multiple areas of frontal and parietal cortex, and that any one of these areas could individually generate a one-dimensional hand trajectory in real time.

This further evidence for narrow spatial localization of function in ECoG suggests that, despite our promising results, the technology for such brain monitoring applications is still evolving. We see large variations in performance between subjects, which may in part be due to the chance nature of the positioning of ECoG electrodes, spaced coarsely 1 cm apart, which may therefore miss or only partially cover relevant highly-localized functional areas. We hypothesize that higher-density ECoG recordings—such as those of Wang et al. [59]—might prove superior in multi-modal decoding performance.

Our study highlighted challenges and areas of improvement for developing real-time decoding systems with ECoG signals. It should be noted that ECoG research is currently primarily feasible only in patients with epilepsy who are candidates for surgical resection and undergo invasive ECoG monitoring. This means that the extent of grid coverage and placement is determined by the clinical needs of the patient rather than the needs of our research. Hence, grid coverage is variable across subjects and often does not cover all the areas of interest for a particular study. Moreover, during the experimentation period, the patients are recovering from a brain surgery, are undergoing withdrawal of their anti-epileptic medications and, consequently, are more susceptible to seizures. These factors affect the physical and cognitive condition and level of cooperation of each subject. In addition, there are variations in the grid laterality, hand dominance, age, location of epileptic focus, across the subjects. Finally, some of the patients may have some degree of functional or structural reorganization based on the etiology of their epilepsy. Overall, these factors can make human ECoG experiments less controlled than non-invasive neuroscientific studies in healthy human subjects or invasive studies in animals. At the same time, in these respects, our present study is similar to the many other ECoG-based studies that have been described in the literature [37, 17, 23, 16, 22, 18, 39]. Despite the issues described above, the results presented in our present and other ECoG studies are usually consistent with expectations based on the general human neuroanatomy or

on results from other imaging modalities.

To mitigate some of the limitations described above, and to acquire data and test a system under such conditions, one needs a computationally inexpensive, portable, robust, and semi-supervised system, with few free parameters. The system should be able to be configured rapidly for each individual subject—perhaps even set up at multiple times of the day—while minimizing inconvenience to the subject and interruptions in their medical treatment. Developing and integrating automated artifact removal, such as for epileptiform spikes, seizures, and other invariant features, may improve the overall real-time signal processing pipeline.

Further research questions that would drive the development of better multi-modal brain-monitoring systems will have to include investigation of how decoding models estimated from synchronous (cue-based) paradigms can best be transferred to asynchronous (non-cue-based) real-time decoding. The ability to transfer decoding models in this way would be very useful for neural engineering applications: a better understanding of the overlap and sharing of structural and functional networks would facilitate development of brain-computer interface technology for the practical replacement of damaged brain functions.

In our study, we describe encouraging progress towards simultaneous real-time decoding of multiple brain systems. The methods and results produced in this study demonstrate for the first time that movement intention and execution can be decoded in real time simultaneously with the decoding of auditory processing. This demonstration advances the understanding of real-time decoding for future performance monitoring and augmentation systems with relevance to a broad range of medical and non-medical applications.

5. Acknowledgements

This work was supported in part by grants from the US Army Research Office [W911NF-08-1-0216 and W911NF-07-1-0415] and the NIH [EB006356 and EB000856]. The authors would like to thank the staff of the AMC Epilepsy Monitoring Unit for their assistance, as well as Griffin Milsap for technical support.

6. References

[1] J R Wolpaw and E W Wolpaw, editors. *Brain-Computer Interfaces: Principles and Practice*. Oxford University Press, 2012.

[2] J R Wolpaw. Brain-computer interfaces as new brain output pathways. *The Journal of Physiology*, 579(Pt 3):613–9, 2007.

[3] J R Wolpaw, N Birbaumer, D J McFarland, G Pfurtscheller, and T M Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–91, 2002.

[4] W Wang, J L Collinger, A D Degenhart, E C Tyler-Kabara, A B Schwartz, D W Moran, D J Weber, B Wodlinger, R K Vinjamuri, R C Ashmore, J W Kelly, and M L Boninger. An electrocorticographic brain interface in an individual with tetraplegia. *PloS One*, 8(2), 2013.

[5] M A L Nicolelis. Brain - machine interfaces to restore motor function and probe neural circuits. *Nature Reviews Neuroscience*, 4:417–422, 2003.

[6] F Mussa-Ivaldi and L E Miller. Brain-machine interfaces: computational demands and clinical needs meet basic neuroscience. *Trends in Neurosciences*, 26(6):329–34, 2003.

[7] J P Donoghue. Connecting cortex to machines: recent advances in brain interfaces. *Nature Neuroscience*, 5 Suppl:1085–8, 2002.

[8] P G Patil and D Turner. The development of brain-machine interface neuroprosthetic devices. *Neurotherapeutics*, 5(1):137–46, 2008.

[9] M Lebedev and M L Nicolelis. Brain-machine interfaces: past, present and future. *Trends in Neurosciences*, 29(9):536–46, 2006.

[10] D J McFarland, A T Lefkowicz, and J R Wolpaw. Design and operation of an EEG-based brain-computer interface with digital signal processing technology. *Behavior Research Methods, Instruments and Computers*, 29(3), 1997.

[11] D J McFarland, W Sarnacki, T M Vaughan, and J R Wolpaw. Brain-computer interface (BCI) operation: signal and noise during early training sessions. *Clinical Neurophysiology*, 116(1):56–62, 2005.

[12] G Schalk, J Kubánek, K J Miller, N R Anderson, E C Leuthardt, J G Ojemann, D Limbrick, D Moran, L Gerhardt, and J R Wolpaw. Decoding two-dimensional movement trajectories using electrocorticographic signals in humans. *Journal of Neural Engineering*, 4(3):264–75, 2007.

[13] J Kubánek, KJ Miller, JG Ojemann, JR Wolpaw, and G Schalk. Decoding flexion of individual fingers using electrocorticographic signals in humans. *Journal of Neural Engineering*, 6:066001, 2009.

[14] T Pistohl, T Ball, A Schulze-Bonhage, A Aertsen, and C Mehring. Prediction of arm movement trajectories from ECoG-recordings in humans. *Journal of Neuroscience Methods*, 167(1):105–14, 2008.

[15] E C Leuthardt, G Schalk, J R Wolpaw, J R Ojemann, and D W Moran. A brain-computer interface using electrocorticographic signals in humans. *Journal of Neural Engineering*, 1(2):63–71, 2004.

[16] Z Wang, Ao Gunduz, P Brunner, A L Ritaccio, Q Ji, and G Schalk. Decoding onset and direction of movements using electrocorticographic (ECoG) signals in humans. *Frontiers in Neuroengineering*, 5(15):1–13, 2012.

[17] N E Crone, D Boatman, B Gordon, and L Hao. Induced electrocorticographic gamma activity during auditory perception. *Clinical Neurophysiology*, 112(4):565–82, 2001.

[18] A Sinai, C W Bowers, C M Crainiceanu, D Boatman, B Gordon, R P Lesser, F A Lenz, and N E Crone. Electrocorticographic high gamma activity versus electrical cortical stimulation mapping of naming. *Brain*, 128:1556–1570, 2005.

[19] E F Chang, J W Rieger, K Johnson, M S Berger, N M Barbaro, and R T Knight. Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience*, 13(11):1428–1433, 2010.

[20] T Ball, A Schulze-Bonhage, A Aertsen, and C Mehring. Differential representation of arm movement direction in relation to cortical anatomy and function. *Journal of Neural Engineering*, 6(1):016006, 2009.

[21] A Flinker, E F Chang, N M Barbaro, M S Berger, and R T Knight. Sub-centimeter language organization in the human temporal lobe. *Brain and Language*, 117(3):103–109, 2011.

[22] C Potes, A Gunduz, P Brunner, and G Schalk. Dynamics of electrocorticographic (ECoG) activity in human temporal and frontal cortical areas during music listening. *NeuroImage*, 61(4):841–8, 2012.

[23] P Brunner, A L Ritaccio, T M Lynch, J F Emrich, J Adam, J C Williams, E J Aarnoutse, N F Ramsey, and E C Leuthardt. A practical procedure for real-time functional mapping of eloquent cortex using electrocorticographic signals in humans. *Epilepsy Behavior*, 15(3):278–286, 2009.

[24] R Flamary and A Rakotomamonjy. Decoding finger movements from ECoG signals using switching linear models. *Frontiers in Neuroscience*, 6(29):1–9, 2012.

[25] E M M Jongen and F T Y Smulders. Sequence effects in a spatial cueing task endogenous orienting is sensitive to orienting in the preceding trial. *Psychological Research*, 71(5):516–523, 2006.

[26] C Mehring, M P Nawrot, S C de Oliveira, E Vaadia, A Schulze-Bonhage, A Aertsen, and T Ball. Comparing information about arm movement direction in single channels of local and epicortical field potentials from monkey and human motor cortex. *Journal of Physiology*, 98(4-6):498–506, 2005.

[27] A Gunduz, P Brunner, A Daitch, E C Leuthardt, A L Ritaccio, B Pesaran, and G Schalk. Decoding covert spatial attention using electrocorticographic (ECoG) signals in humans. *NeuroImage*, 60(4):2285–2293, 2012.

[28] J Kubánek, P Brunner, A Gunduz, D Poeppel, and G Schalk. The tracking of speech envelope in the human cortex. *PloS one*, 8(1):e53398, 2013.

[29] B R Townsend, E Subasi, and H Scherberger. Grasp movement decoding from premotor and parietal cortex. *Journal of Neuroscience*, 31(40):14386–98, 2011.

[30] W Wu, M J Black, D Mumford, Y Gao, E Bienenstock, and J P Donoghue. Modeling and decoding motor cortical activity using a switching kalman filter. *IEEE Transactions on Biomedical Engineering*, 51(6):933–42, 2004.

[31] W Wu and N G Hatsopoulos. Real-time decoding of nonstationary neural activity in motor cortex. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 16(3):213–22, 2008.

[32] J M Carmena, M Lebedev, R E Crist, J E O'Doherty, D M Santucci, D F Dimitrov, P G Patil, C S Henriquez, and M L Nicolelis. Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biology*, 1(2):193–208, 2003.

[33] N Hatsopoulos, J Joshi, and J G O'Leary. Decoding continuous and discrete motor behaviors using motor and premotor cortical ensembles. *Journal of Neurophysiology*, 92(2):1165–74, 2004.

[34] J Wessberg, C R Stambaugh, J D Kralik, P D Beck, M Laubach, J K Chapin, J Kim, S J Biggs, M Srinivasan, and M Nicolelis. Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. *Nature*, 408(6810):361–5, 2000.

[35] Z C Chao, Y Nagasaka, and N Fujii. Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. *Frontiers in Neuroengineering*, 3:1–4, 2010.

[36] K J Miller, Den Nijs, Shenoy M., Miller P., J W, R P Rao, and J G Ojemann. Real-time functional brain mapping using electrocorticography. *Neuroimage*, 37:504–507, 2007.

[37] K J Miller, E C Leuthardt, G Schalk, R P N Rao, N R Anderson, D W Moran, J W Miller, and J G Ojemann. Spectral changes in cortical surface potentials during motor movement. *The Journal of Neuroscience*, 27(9):2424–32, 2007.

[38] N E Crone, A Sinai, and A Korzeniewska. High-frequency gamma oscillations and human brain mapping with electrocorticography. *Progress in Brain Research*, 159:275–295, 2006.

[39] A Sinai, N E Crone, H M Wied, P J Franaszczuk, D Miglioretti, and D Boatman-Reich. Intracranial mapping of auditory perception: event-related responses and electrocortical stimulation. *Clinical Neurophysiology*, 120(1):140–9, 2009.

[40] N J Hill, D Gupta, P Brunner, A Gunduz, M A Adamo, A Ritaccio, and G Schalk. Recording human electrocorticographic (ECoG) signals for neuroscientific research and real-time functional cortical mapping. *Journal of Visualized Experiments*, (64):1–5, 2012.

[41] G Schalk, D J McFarland, T Hinterberger, N Birbaumer, and J R Wolpaw. BCI2000: a general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, 51(6):1034–43, 2004.

[42] G Schalk and J Mellinger. *A Practical Guide to Brain-Computer Interfacing with* BCI*2000*. Springer, London, UK, 1st edition, 2010.

[43] J L Lancaster, L H Rainey, J L Summerlin, C S Freitas, P T Fox, A C Evans, A W Toga, and J C M. Automated labeling of the human brain : A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping*, (5):238–242, 1997.

[44] J L Lancaster, M G Woldorff, L M Parsons, M Liotti, C S Freitas, L Rainey, P V Kochunov, D Nickerson, S A Mikiten, and P T Fox. Automated talairach atlas labels for functional brain

mapping. *Human Brain Mapping*, 10(3):120 –131, 2000.

[45] J P Burg. Maximum entropy spectral analysis. In *37th Meet. Soc. Exploration Geophysicists*, Oklahoma, 1967.

[46] S Lemm, B Blankertz, T Dickhaus, and K-R Müller. Introduction to machine learning for brain imaging. *NeuroImage*, 56(2):387–99, 2011.

[47] P Shenoy, K J Miller, J G Ojemann, and R P N Rao. Generalized features for electrocorticographic BCIs. *IEEE Transactions on Biomedical Engineering*, 55(1):273–80, 2008.

[48] M Velliste, S Perel, M C Spalding, A S Whitford, and A B Schwartz. Cortical control of a prosthetic arm for self-feeding. *Nature*, 453(7198):1098–101, 2008.

[49] L R Hochberg, D Bacher, B Jarosiewicz, N Y Masse, J D Simeral, J Vogel, S Haddadin, J Liu, S S Cash, P van der Smagt, and J P Donoghue. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature*, 485(7398):372–5, 2012.

[50] J Viventi, D-H Kim, L Vigeland, E S Frechette, J Blanco, Y-S Kim, A E Avrin, V R Tiruvadi, S-W Hwang, A C Vanleer, D F Wulsin, K Davis, C E Gelber, L Palmer, J Van der Spiegel, J Wu, J Xiao, Y Huang, D Contreras, J Rogers, and B Litt. Flexible, foldable, actively multiplexed, high-density electrode array for mapping brain activity in vivo. *Nature Neuroscience*, 14(12):1599–605, 2011.

[51] J Viventi, Dae-H Kim, J D Moss, Y-S Kim, J Blanco, N Annetta, A Hicks, J Xiao, Y Huang, D J Callans, J Rogers, and B Litt. A conformal, bio-interfaced class of silicon electronics for mapping cardiac electrophysiology. *Science Translational Medicine*, 2(24), 2010.

[52] C Henle, M Raab, J G Cordeiro, S Doostkam, T Stieglitz, and J Rickert. First long term in vivo study on subdurally implanted micro-ECoG electrodes , manufactured with a novel laser technology. *Biomedical Microdevices*, 13(1):59–68, 2011.

[53] D Borton, M Yin, J Aceros, and A Nurmikko. An implantable wireless neural interface for recording cortical circuit dynamics in moving primates. *Journal of Neural Engineering*, 10(2), 2013.

[54] G S. Anderson and R R. Harrison. Wireless integrated circuit for the acquisition of electrocorticogram signals. *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, pages 2952–2955, 2010.

[55] C Chestek, V Gilja, P Nuyujukian, R J Kier, F Solzbacher, S I Ryu, R R Harrison, and K V Shenoy. HermesC: low-power wireless neural recording system for freely moving primates. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(4):330–8, 2009.

[56] D Fan, D Rich, T Holtzman, P Ruther, J W Dalley, A Lopez, M Rossi, J W Barter, D Salas-Meza, S Herwik, T Holzhammer, J Morizio, and H H Yin. A wireless multi-channel recording system for freely behaving mice and rats. *PloS One*, 6(7), 2011.

[57] E Tolstosheeva, V Gordillo-González, T Hertzberg, L Kempen, I Michels, A Kreiter, and W Lang. A novel flex-rigid and soft-release ECoG array. In *IEEE Engineering in Medicine and Biology Society Conference*, pages 2973–6, 2011.

[58] P Brunner, L Bianchi, C Guger, F Cincotti, and G Schalk. Current trends in hardware and software for brain-computer interfaces (BCI). *Journal of Neural Engineering*, 8(2):1–7, 2011.

[59] W Wang, A D Degenhart, J L Collinger, R Vinjamuri, G P Sudre, P D Adelson, D L Holder, E C Leuthardt, D W Moran, M L Boninger, A B Schwatrz, D J Crammond, E C Tyler-Kabara, and D J Weber. Human motor cortical activity recorded with micro-ECoG electrodes during individual finger movements. In *IEEE Engineering in Medicine and Biology Magazine*, number 1, pages 586–589, 2009.