NAIST-IS-DD1161012

Doctoral Dissertation

Neural Decoding of Electrocorticographic Signals

Kei Majima

February 6, 2014

Department of Bioinformatics and Genomics
Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of SCIENCE

Kei Majima

Thesis Committee:

Professor Kazushi Ikeda	(Supervisor)
Professor Kenji Sugimoto	(Co-supervisor)
Professor Mitsuo Kawato	(Co-supervisor)
Professor Yukiyasu Kamitani	(Co-supervisor)

Neural Decoding of Electrocorticographic Signals*

Kei Majima

Abstract

Over the last decade, neural decoding technology based on machine learning has enabled us to extract fine information on visual experiences and motor commands from measured neural signals, which is becoming a powerful tool for investigating neural representations and generating commands for controlling brain–machine interfaces (BMIs). To extract information with high predictive performance, neural recording methods that provide high spatiotemporal resolution and signal stability are required. One candidate is electrocorticogram (ECoG), which measures population activity of neurons with electrodes placed on the surface of the brain. Here, with the aim of high-performance decoding with ECoG, we tested the utility of ECoG systems in animal and human studies, and improved techniques to extract information from ECoG data. As the first contribution of this thesis, the signal stability of ECoG responses recorded via a newly developed high-density mesh electrode array was tested. Collaborators applied it to the visual cortex in rats and this thesis demonstrates above-chance, generalized decoding performance for simple visual stimulation, using six hours of continuous data (chapter 2). Second, by applying decoding analysis to simultaneously recorded ECoG, LFP, and MUA signals from the monkey IT cortex, extractable information on visually presented objects was compared. The resultant decoding performance with ECoG was high and comparable to LFP and MUA (chapter 3). Next, ECoG was used to investigate how face-selective regions and written word-selective regions are distributed on the human cortex, which is considered a challenging task with fMRI. Results reveal that there exist multiple, separate face- and written word-selective regions in the human cortex (chapter 4). Finally, using ECoG responses from human patients when they viewed

*Doctoral Dissertation, Department of Bioinformatics and Genomics, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-DD1161012, February 6, 2014.

objects, efficient input signal features for decoding analysis were explored. Spectral powers, phases, and temporal correlations of ECoG signals were used as input features, and the decoding performances were compared. Results show the performance using temporal correlations between ECoG electrodes is higher than using spectral powers and phases in individual electrodes (chapter 5). Those results suggest that combining ECoG recordings and neural decoding techniques is a powerful approach for extracting neural information, and we can considerably improve a decoder's predictive performance by using signal features that take into account fine temporal patterns in ECoG signals.

Keywords:

Neural decoding, ECoG

Contents

List of Figures

List of Tables

1. General Introduction

The recent advancement of neural decoding technology has enabled us to extract fine information from measured neural signals, which provides an important basis for understanding visual, motor, and cognitive representations in the brain. Originally, neural decoding techniques were developed for spike data obtained by single-unit recording (Bialek et al., 1991; Dayan and Abbott, 2001; Stanley, 2013). Recently, by analyzing activity patterns over multiple voxels, neural decoding techniques have been applied to fMRI signals and are becoming a powerful way to reveal neural representations in humans. Haxby et al. (2001) demonstrated that broadly distributed fMRI activity patterns in the temporal cortex can be used to indentify the category of presented visual objects, even though they cannot be easily distinguished by conventional univariate analyses. In more recent studies involving more sophisticated machine-learning algorithms, it has been shown that information on low-level visual features such as color, orientation, contrast, and information on subjective experiences such as imagined objects and dreams can be reliably predicted from single-trial fMRI signals (Harrison and Tong, 2009; Horikawa et al., 2013; Kamitani and Tong, 2005,2006; Miyawaki et al., 2008; Naselaris et al., 2009; Nishimoto et al., 2011; Reddy et al., 2010; Yamamura et al., 2009).

Also, neural decoding technology has been used to generate commands for brain–machine interfaces (BMIs), such as a prosthetic device for paralyzed people (Lebedev and Nicolelis, 2006; Leuthardt et al., 2006). Previous studies have shown that intended hand motions or trajectories can be estimated from recorded neuronal ensemble activity (Donoghue, 2002; Georgopoulos et al., 1986; Hochberg et al., 2006; Serruya et al., 2002; Truccolo et al., 2008; Wessberg et al., 2000). Therefore, neural decoding systems with high-predictive performance have been required.

To extract information with high predictive performance over a long term, neural recording methods that provide high spatiotemporal resolution and stability are required. One promising tool is electrocorticography (ECoG). ECoG is a methodology for mapping (recording) neural electric signals from a wide cortical region with high signal fidelity by placing electrodes on the surface of the brain (Fig. 1.1). ECoG was developed originally as a clinical tool for monitoring and mapping seizure foci during neurosurgery and, thanks to the recent advancement of electrodes and their probes, is now recognized as a potent methodology for basic neurophysiological investigations

(Fisch et al., 2009; Liu et al., 2009; Tsuchiya et al., 2008) and BMI applications (Chao et al., 2010; Leuthardt et al., 2004; Pistohl et al., 2008; Schalk et al., 2008; Shimoda et al., 2012; Yanagisawa et al., 2009).

Compared with other neuroimaging techniques such as functional magnetic resonance imaging (fMRI), scalp electroencephalography (scalp EEG), and magnetoencephalography (MEG), ECoG provides a balanced tradeoff between spatial and temporal resolutions. Moderately good spatial resolution can be attained using fMRI, yet its temporal resolution is limited to the timescale of seconds, which is much lower than ECoG. On the other hand, scalp EEG and MEG offer millisecond temporal resolution, which is as good as ECoG, but their spatial resolutions (on the order of several centimeters) are lower than ECoG. Although direct single-unit recordings offer the best possible spatio-temporal resolution in principle, this technique suffers from an extremely narrow anatomical field-of-view, together with very rare opportunities to obtain such recordings in humans.

Here, toward the construction of high-performance decoding systems with ECoG, we made a new ECoG electrode array, tested the utility of ECoG systems with animals and human patients, and developed an efficient way of decoder construction for ECoG data.

This research was conducted at ATR Computational Neuroscience Laboratories, Niigata University School of Medicine, The University of Tokyo Hospital, and Nishi-Niigata Chuo National Hospital (see ***Introduction*** in each chapter for details). The development of ECoG electrode arrays and all animal experiments were performed by researchers at Niigata University School of Medicine. All experiments with human patients were performed by researchers at The University of Tokyo Hospital and Nishi-Niigata Chuo National Hospital. The author of this thesis conducted all decoding analyses in those animal and human studies, and mainly conducted the analyses described in chapter 5.

In chapter 2, we developed a high-density electrode array and tested its signal stability by neural decoding. Recently, high-density ECoG electrode arrays have been developed and are becoming a promising tool for electrophysiological mapping with a precision of millimeters. By utilizing new technology developed by Takeuchi et al. (2005), we developed a high-density electrode array by placing electrodes on a flexible, thin mesh

and tested its ECoG responses with rats. Here, to verify the signal stability over time, a decoder's generalization performance was examined and kept above the chance level for six hours of data.

In chapter 3, the decoding performances obtained with simultaneously recorded ECoG, local field potential (LFP), and multi-unit activity (MUA) signals were compared. Direct comparison of these signals is one of general interest. However, simultaneous measurement of LFP and MUA signals with a standard ECoG array is a challenging task, due to their configurations. To address this issue, we developed an ECoG electrode-mesh and succeeded in simultaneously recording ECoG, LFP, and MUA signals from the monkey IT cortex. In the experiments, images of objects were presented to the monkeys and ECoG, LFP, and MUA responses were simultaneously recorded. Here, we report the first direct comparison of decoding performance between these recording modalities.

In chapter 4, by analyzing ECoG responses from human patients, we demonstrate that ECoG is a useful recording method to localize category-responsive regions in the vicinity of the bottom of the brain, which is considered a challenging task using fMRI. Previous studies have shown that there exist brain regions that mainly process information on specific object categories such as faces and written words. How those face- and written word-responsive regions are distributed in the temporal cortex is one of the most important questions in neuroscience. However, it is challenging because fMRI responses in the vicinity of the bottom of the brain suffer from artifacts by considerable inhomogeneity of the blood oxygen level–dependent signals (Ojemann et al., 1997; Winawer et al., 2010). To avoid this, we recorded ECoG responses from human patients and localized brain regions responsive to face stimuli and written word stimuli through a standard electrode-wise analysis. To confirm the results, we also applied decoding analysis. The results from these analyses suggest that there exist multiple, separated face-responsive regions and written word-responsive regions in the cortex.

In chapter 5, using ECoG responses from the human temporal cortex, a decoder's predictive performance was improved by using temporal correlations of ECoG signals between electrodes. In previous studies, spectral powers and response amplitudes were often highlighted and used as input features for decoding analysis. However, previous studies have shown that synchronization and phase differences across brain regions carry a significant amount of information on low-level visual features. If such a coding

is adopted in the representation of visual objects, the use of temporal correlations of ECoG signals would improve the decoding performance. The theory that information on objects is encoded in phase differences across brain regions is also supported by simulation analysis.

This thesis shows that the combination of ECoG recordings and neural decoding techniques is a powerful approach for extracting neural information, and we can considerably improve a decoder's predictive performance by using signal features that take into account fine temporal patterns in ECoG signals.

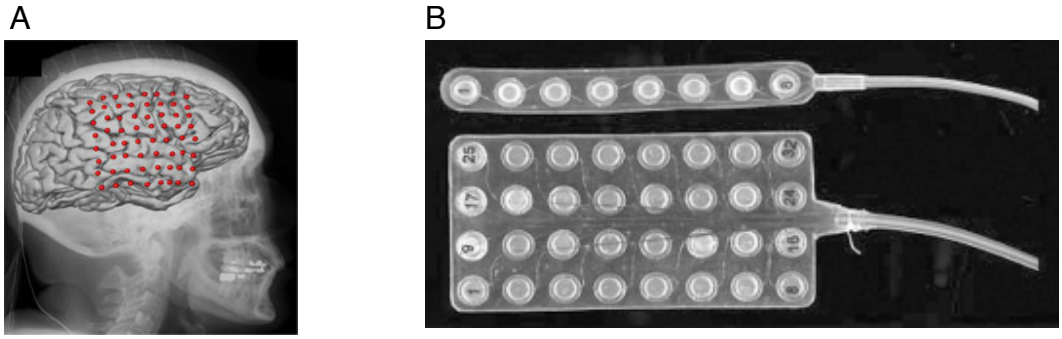

Figure 1.1. Electrocortocogram (ECoG). ***A,*** An x-ray picture of a patient's brain with an overlay indicating the locations of ECoG electrodes. ***B,*** Examples of ECoG electrode arrays.

2. ECoG responses recorded with a flexible electrode-mesh and its signal stability

2.1 Introduction

Recently, new, high-density ECoG electrode arrays are being developed that show promise as a tool for electrophysiological mapping with a precision of millimeters (Hollenberg et al., 2006; Rubehn et al., 2009; Yeager et al., 2008). Toda et al. (2011) placed electrodes on a thin, flexible mesh by utilizing a new technology (Takeuchi et al., 2005), and developed a high-density ECoG electrode array (Fig 2.1*A*). Thanks to this thin, flexible electrode mesh, their ECoG probe was able to better fit to the curved surface of the brain and is considered less invasive than conventional clinical probes.

When applying such new electrodes to *in vivo* experiments, whether the recording system can provide stable signals across multiple sessions or days is one of the main concerns in electrophysiological experiments. To collect enough data to investigate neural representations, an experiment that consists of multiple sessions over hours is often required. Also, to make brain–machine interface (BMI) systems practical in real life, the long-term stability of input signals is required. Previous studies have shown that ECoG can provide signals stable across even several months (Chao et al., 2010; Rubehn et al., 2009; Shimoda et al., 2012). However, it is possible that signal stability is sensitive to recording environments and affected by conditions.

To evaluate the signal stability of ECoG reponses obtained with the electrode mesh, they applied it to the visual cortex in rats (Fig. 2.1***B,C***). In their experiments, simple light stimulation was given to one of the eyes. They recorded ECoG responses to it (Fig 2.2) and showed that the trial-averaged signals were stable over six hours (Fig 2.3).

Here, we performed decoding analysis to examine whether not only the trial-averaged waveforms, but also extractable information on visual stimuli is stable across sessions. A statistical classifier was constructed to predict which eye was

stimulated on a single-trial basis, and the decoder's predictive performance was tested with ECoG data recorded several hours after the initial condition.

From the analyses conducted here, it is shown that Toda et al.'s (2011) ECoG preparations could provide stable signals for at least six hours. This work was published as part of Toda et al. (2011) and the author of this thesis performed the decoding analysis in it.

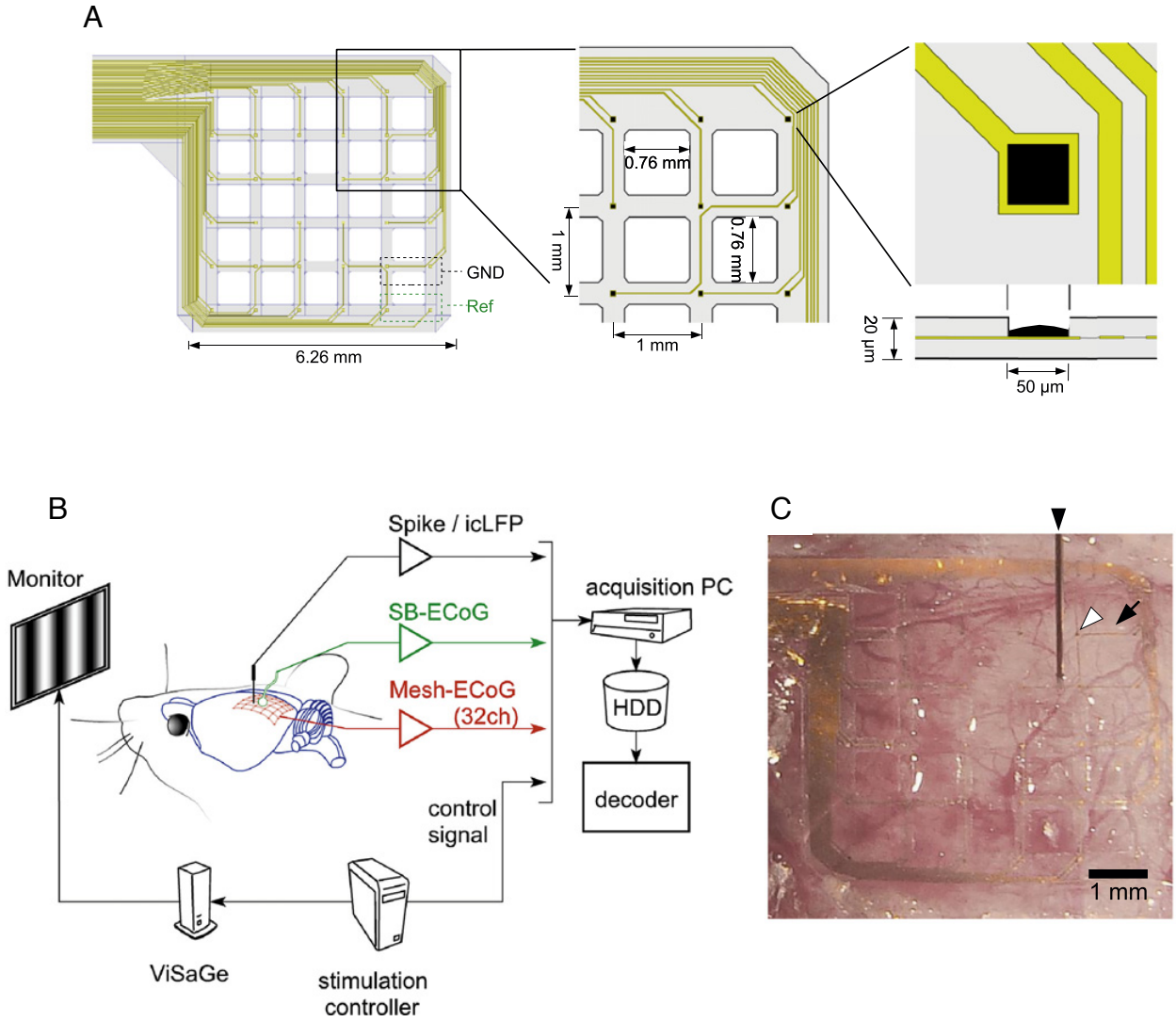

Figure 2.1. Flexible electrode-mesh: structure and *in vivo* application. ***A,*** Structure of an electrode-mesh. Thirty-six gold electrodes (50×50 μm) were arranged in a 6×6 array with 1-mm inter-electrode intervals. ***B,*** Experimental setup for in vivo application of the electrode-mesh. During visual stimulation, multichannel ECoGs via an electrode mesh (mesh ECoG), single-channel ECoG via a silver-ball ECoG electrode (SB) and spikes or intracortical LFPs (icLFP) via a tungsten microelectrode were simultaneously recorded from the rat visual cortex. ***C,*** Epidural placement of an electrode-mesh and a tungsten microelectrode. White arrowhead: one of the ECoG electrodes.

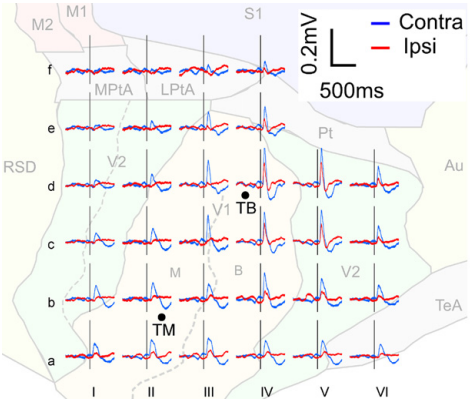

Figure 2.2. Visually evoked ECoG signals from an electrode mesh. Multichannel ECoG responses to Contra (blue traces) and Ipsi (red traces) stimulation are superimposed on the cortical map. Vertical black line in each panel indicates the stimulus onset. Numbers of averaged trials were 40 and 39 for the Contra and Ipsi conditions, respectively.

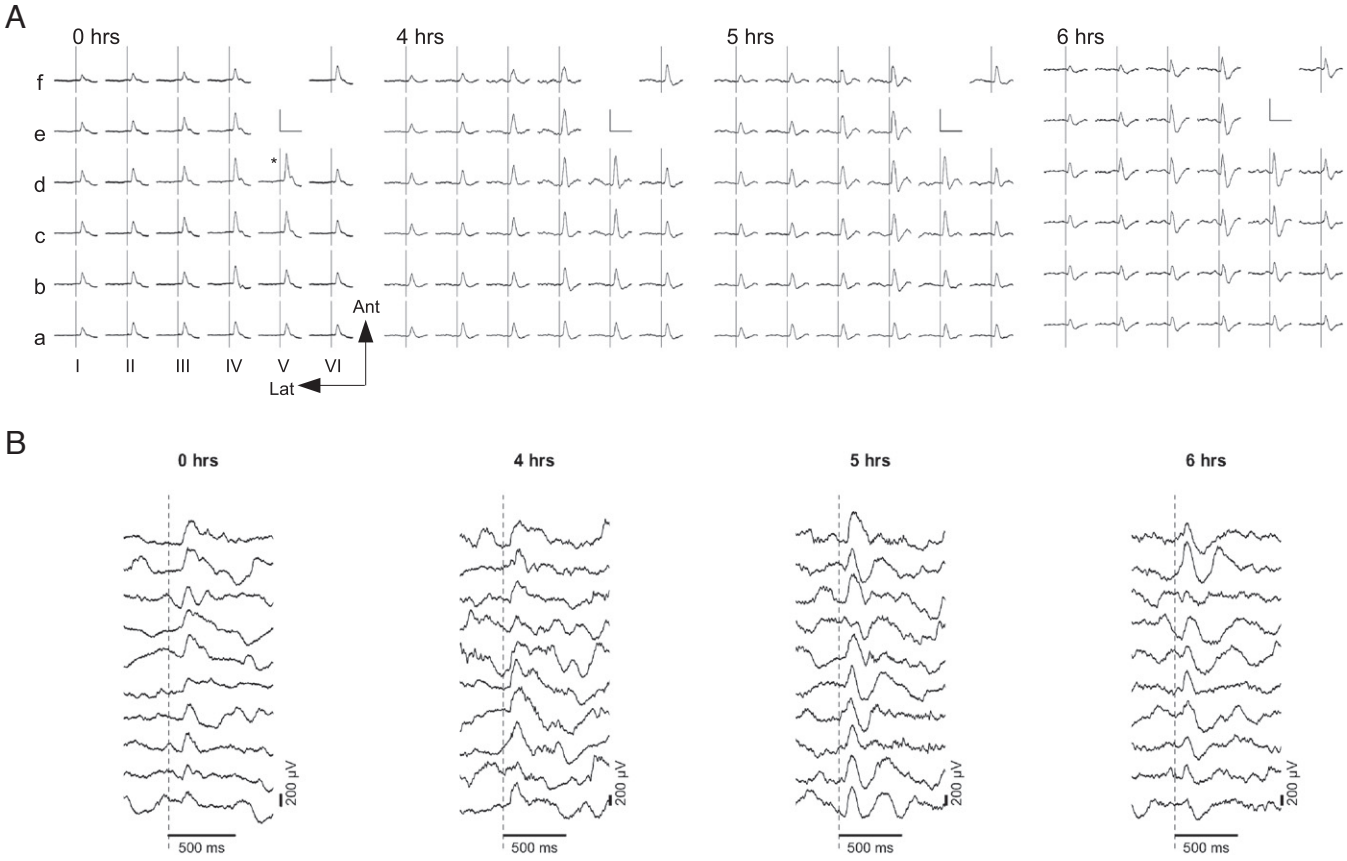

Figure 2.3. Across- and within-session stabilities of multichannel ECoG signals from the electrode mesh. ***A,*** Trial-averaged visually-evoked responses recorded 4, 5, and 6 hours following the initial session (0 hours). The inset indicates the simultaneously recorded icLFP for reference. Scale bars: 0.2 mV and 500 ms. Recordings were reliably maintained for 6 hours, although slight variations in waveforms were similarly found in simultaneously recorded ECoG and icLFP data. Ant: anterior (rostral), Lat: lateral. ***B,*** Single-trial visually evoked responses across sessions. Visually evoked responses in ten consecutive trials were shown from top to bottom in each column. Signals were recorded at the electrode Vd (marked by asterisk) in the same sessions as *A*.

2.2 Materials and Methods

2.2.1 Experimental data

Data were recorded in Niigata University School of Medicine under the supervision of Dr. Toda, Dr. Sawahata, and Dr. Hasegawa. Their recording conditions were briefly described in the following sections (see Toda et al. (2011) for experimental details).

Structure of the electrode-mesh
An electrode-mesh (Fig. 2.1*A*) was designed to contain 32-channel recording electrodes plus 2 reference and 2 ground electrodes. The inter-electrode distance was 1 mm. The whole mesh size was 6 × 6 mm. Detailed fabrication procedures for the electrodes were described in Takeuchi et al. (2005).

Animal preparations
A schematic illustration of the recording setup is shown in Fig. 2.1*B*. Data were obtained from six hemispheres in three adult rats (Long–Evans, 300–400 g). The rats were anesthetized by intraperitoneal injection of sodium pentobarbitulate. The anesthetized rats were restrained to the stereotaxic apparatus (SR-5, Narishige, Tokyo, Japan) with their head fixed. Part of the cranium was removed to place the electrode-mesh on the dura mater and to allow tungsten microelectrode penetrations into the cortex (Fig. 2.1*C*). The recording areas corresponded to V1 and the adjacent secondary visual cortex (V2) except for the posterior tip of V1 (Fig. 2.2) in the unrolled cortical map reconstructed from the rat brain (Paxinos and Watson, 2007). All the experimental procedures were approved by the Niigata University Animal Experiment Committee and conformed to guidelines of the animal welfare laws in Japan and NIH Guide for the Care and Use of Laboratory Animals.

Visual stimulation
Each anesthetized animal was placed 28.6 cm in front of a 19-inch display monitor (EIZO T561, Nanao, Ishikawa, Japan). Visual stimuli were presented to the full field of the display. Visual stimuli were moving sinusoidal gratings. Spatial

(0.04–0.15 cycle/deg) and temporal (3–7 Hz) frequencies of the gratings were optimized according to Girman et al. (1999). Stimulus duration was 500 ms and the inter-trial interval was pseudo-randomized between 1.5 and 2.0 sec. In the contralateral-eye stimulation (Contra) or ipsilateral-eye stimulation (Ipsi) conditions, the eye ipsilateral or contralateral to the recorded hemisphere, respectively, was covered with an eye patch. Stimulation was performed in separate blocks consisting of 40–200 trials.

Data acquisition
32-channel ECoG signals were sampled at 1kHz, then differentially amplified and bandpassed (300 Hz and 0.7 Hz high- and low-cutoff frequencies). The signal references for multichannel ECoGs were connected to the reference on the electrode mesh.

2.2.2 Decoding analysis

Input features were extracted from ECoG signals during a period from -25 ms to 325 ms, relative to the stimulus onset for each trial. The amplitudes for each channel were sampled by a 50-ms time window shifted by 50 ms, and the power spectrum in each time window was calculated. The powers of each frequency, time window and channel were used as input features (a feature vector) for classification analysis (7 time windows $\times$ 5 frequency bands $\times$ 32 channels; total of 1120 features in each data sample for the electrode mesh).

The feature vector obtained from each trial was labeled with the stimulated eye (left eye or right eye). The dimensionality of the feature vectors was reduced by selecting informative features based on a univariate analysis (t-statistics) applied to a training data set. The features were ranked by the t-value that indicated the differential responses of the two stimulus conditions (left eye or right eye) and the top N features (n = 1, 5, 10, 25, 50, 75, 100, 150, and 200) were selected. The N-dimensional feature vectors served as inputs to the decoding model.

Decoding performance was evaluated through two types of analysis. First, to evaluate the overall performance of each rat/hemisphere avoiding the circularity

that can arise from data selection (Kriegeskorte et al., 2009), we performed a cross-validation analysis using the entire data set. The data were divided into 10 subgroups, 9 of which were used to train a classifier, and the remaining subgroup was used for evaluating the trained classifier. This procedure was repeated until the samples from all 10 subgroups were tested (10-fold cross-validation) and the percentage of correct classification was calculated.

Second, to evaluate the stability of decoding with multichannel ECoG responses, a classifier was trained using the data measured during the first hour of the experiment and tested on the blocks of data measured in the following 1–6 hours (1 hour in 1 hemisphere, 3 hours in 3 hemispheres, 4 hours for 1 hemisphere and 6 hours for 1 hemisphere). The selected top 100 features were used as inputs. For the first hour, cross-validation classification accuracies were calculated. For each of the following hours, generalization accuracies were calculated. The binomial test was applied to examine whether the frequency of correct classification exceeded the chance level.

2.3 Results

Toda et al. (2011) succeeded in recording 32-channel ECoG responses with a novel flexible ECoG electrode array (electrode mesh) from the early visual cortex in rats (Fig. 2.1). They have shown that the trial-averaged ECoG responses had clear selectivity for two different visual stimulation conditions (Contra-eye stimulation and Ipsi-eye stimulation) (Fig. 2.2). Also, they compared ECoG responses across sessions and showed that visually-evoked responses could be observed six hours after the beginning of the recording (Fig 2.3). In the following, we confirm that those ECoG responses could be used for decoding analysis and a decoder's predictive performance across sessions kept above the chance level.

We examined whether the stimulation condition (i.e., left eye or right eye) could be reliably decoded from single-trial ECoG signals by a statistical classifier (linear support vector machine, SVM). Correct classification rate was around seventy percent with top-ranked features derived from the ECoG signals (see ***Materials and Methods***), and increased with the number of features, approaching ninety percent at 50–100 features (Fig. 2.4*A*).

Consistent with across-session stability of the waveforms of the mesh ECoG signals (Fig. 2.3), generalization accuracies of the classifier across sessions (training the decoder on the first session, and testing it on subsequent sessions) were statistically significantly above the random chance level for 6 hours through the test period (Fig. 2.4*B*).

Stimulus-locked topographical weight maps (Fig. 2.5) show spatiotemporal profiles of effective information used in the decoding-based analysis. Channels with high weight values were mainly found 75–175 ms following stimulation onset, and tended to form a few clusters.

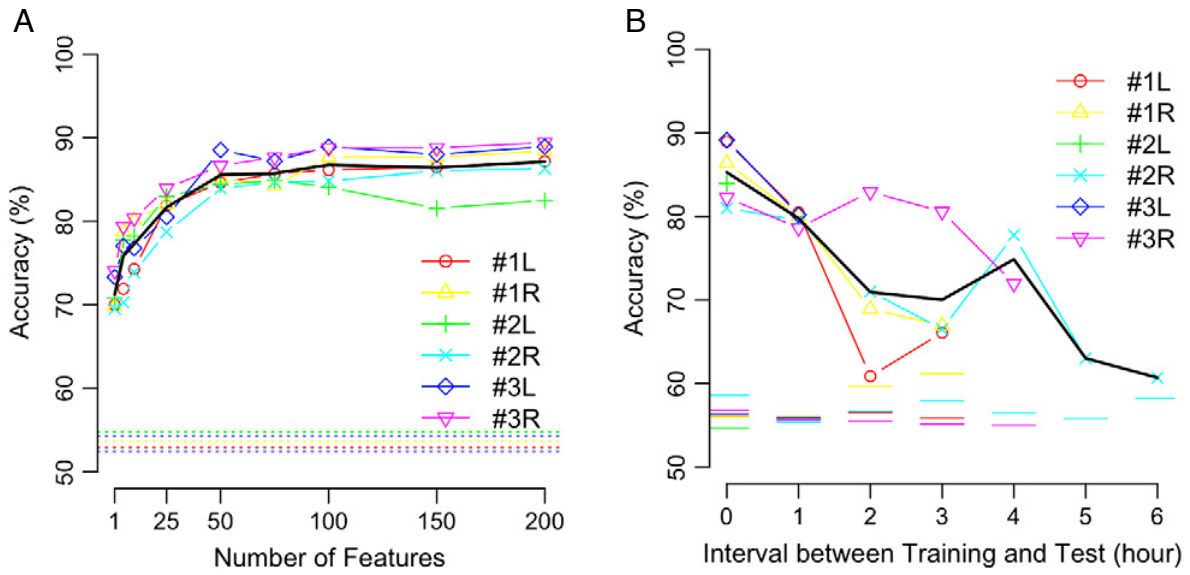

Figure 2.4. Single-trial decoding accuracy by ECoG signals from the electrode-mesh. ***A,*** Decoding accuracies on single-trial mesh ECoG signals plotted against the number of features used in the analysis. Different colors/symbols represent different hemispheres of the rat brains. Colored, dotted lines indicate 1% significance levels for each hemisphere of the same color. Accuracy increased with the number of features, approaching 90% at 50–100 features. ***B,*** Generalization accuracy of the classifier for hourly repeated recording sessions with top-100 features from the mesh ECoG signals. The abscissa indicates time after the initial session wherein the classifier had been trained. The performance was above the 1% significance level (solid line of the same color) for all sessions and hemispheres.

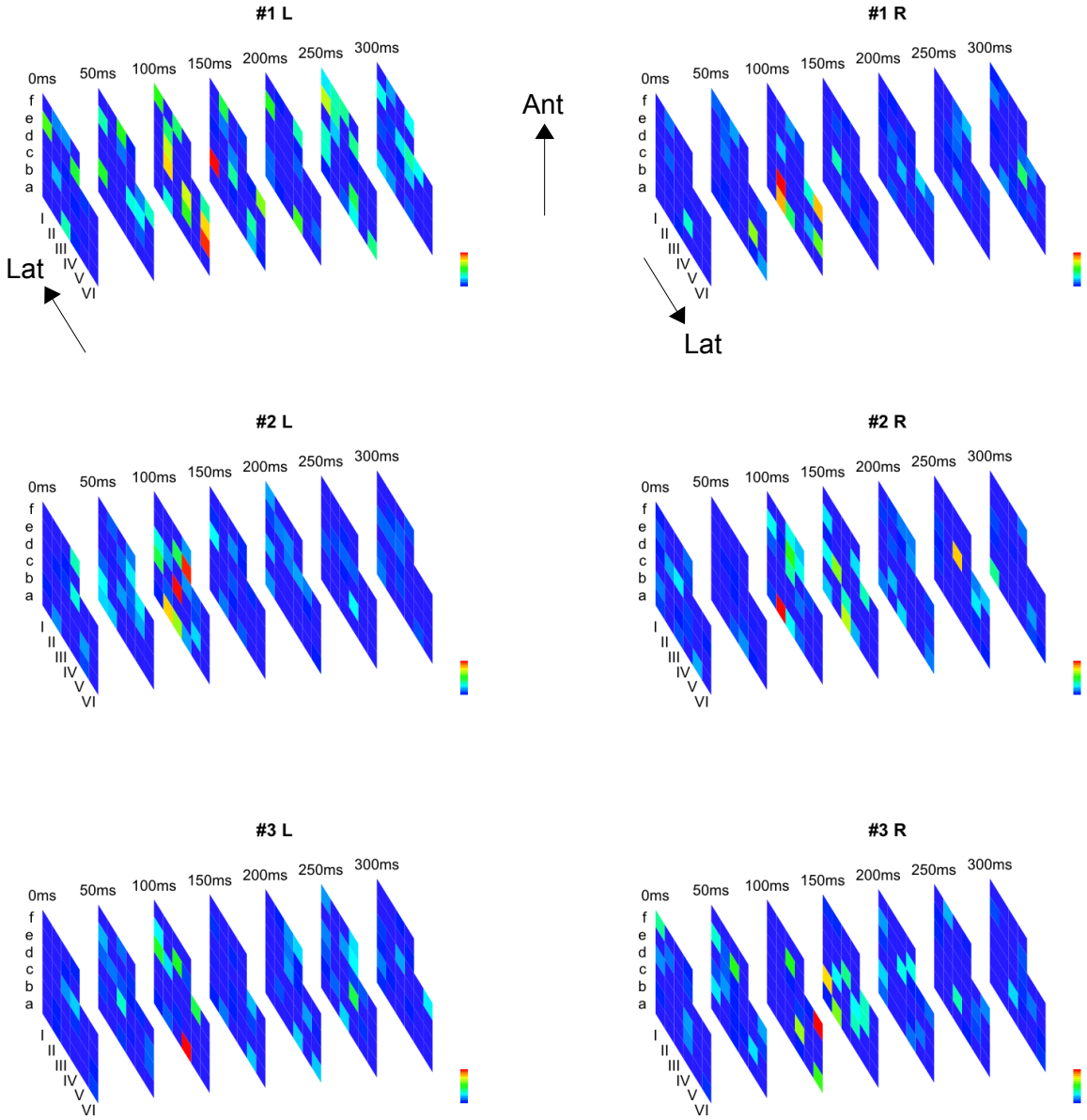

Figure 2.5. Weight maps for the decoders in individual rat/hemisphere. Normalized mean absolute weights (averaged across 5 frequency bands) are shown by color for each channel and time window. In this analysis, top-100 features were selected, and an SVM model was trained using all data sets. The weights for non-selected features were set to zero, and the weights for selected features were calculated from the trained SVM model. The weights for non-selected features were assumed to be zero. Ant: anterior (rostral), Lat: lateral.

2.4 Discussion

We found that single-trial ECoG data contained information to reliably classify the stimulated eye with an accuracy approaching 90% (Fig 2.4*A*). Such high single-trial decoding performance complements the results of other analyses based on averaged and channel-by-channel data (Fig 2.2), and provides evidence for the feasibility of using an electrode mesh as a signal input device to BMI systems.

Also, we found that the generalization accuracies of the classifiers across sessions stayed above the chance level over six hours of data (Fig 2.4*B*). Consistent with this result, long-term decoding stability for months was reported with motor-related ECoG signals in monkeys (Chao et al., 2010). In addition, the electrode mesh with a thickness of 20 μm is less invasive than conventional clinical ECoG sheets with a thickness of approximately 1 mm that might cause mass effects when inserted in the cranium. This structure might contribute to the resulting stability.

Although further clinical tests are needed to establish long-term stability for months and years, the results shown in Toda et al. (2011) and here support the view that the electrode mesh is one of the promising candidates for next-generation input devices for BMI (Lebedev and Nicolelis, 2006; Leuthardt et al., 2006).

3. Direct comparison of ECoG, LFP, and MUA

3.1 Introduction

Direct comparison of ECoG signals and neuronal signals provided by other recording methods is of great interest in neuroscience. However, simultaneous measurement of ECoG, LFP, and MUA signals is a challenging task because it is virtually impossible to insert a microelectrode for LFP and MUA recordings into the brain surface through a standard ECoG electrode sheet from the outside.

By introducing an ECoG electrode array that has a mesh structure, a few previous studies have conducted simultaneous recording of ECoG, LFP, and MUA, and directly compared them (Toda et al., 2011; Watanabe et al., 2012). In those studies, ECoG electrodes were arranged on a mesh and microelectrodes were inserted into the brain surface through the holes in the mesh. However, their comparison was not fair in terms of the number of electrodes and their spatial coverage.

To do a fair comparison, Miyakawa et al. (2012) arranged 64 ECoG electrodes and 64 microelectrodes with almost the same coverage into the monkey inferiortemporal cortex. (Fig 3.1). They recorded ECoG, LFP, and MUA responses simultaneously when the monkeys viewed images of objects (Fig 3.2).

Here, to compare extractable information on presented object categories between those recording methods, decoding analysis was applied to each of the recording methods, and the resultant performances were compared.

From the analyses conducted in this chapter, it was shown that ECoG can provide good signal fidelity, comparable with LFP and MUA for extracting neural information.

This work has been published as part of Miyakawa et al. (2011), and will be published as part of Miyakawa et al. (in preparation). The author of this thesis performed the decoding analyses in these publications.

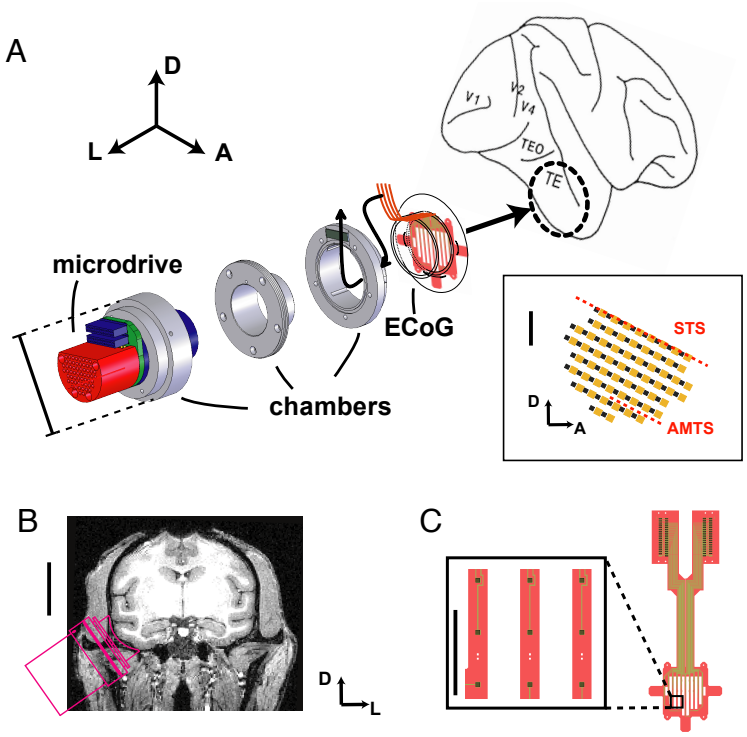

Figure 3.1. Recording setup for simultaneous ECoG, MUA, and LFP recording in macaque IT cortex. ***A,*** Experimental setup for application of an electrode mesh. The ECoG electrode mesh was attached to the titanium chambers and the microdrive system with microelectrodes. (Inset) Schematic drawing of ECoG (yellow) and microelectrode (black) spatial configuration as placed on the cortex. Electrodes were placed on area TE of the IT cortex, covering the IT gyrus and extending slightly below AMTS. ***B,*** Anatomical image with electrode assembly. An MRI image was superimposed with the electrode assembly. The assembly was placed to cover the IT gyrus ***C,*** 2D schematic drawing of the original ECoG electrode array. The ECoG array was made using MEMS technology. Yellow lines are gold wires. Brown squares (blown up image) are ECoG contacts.

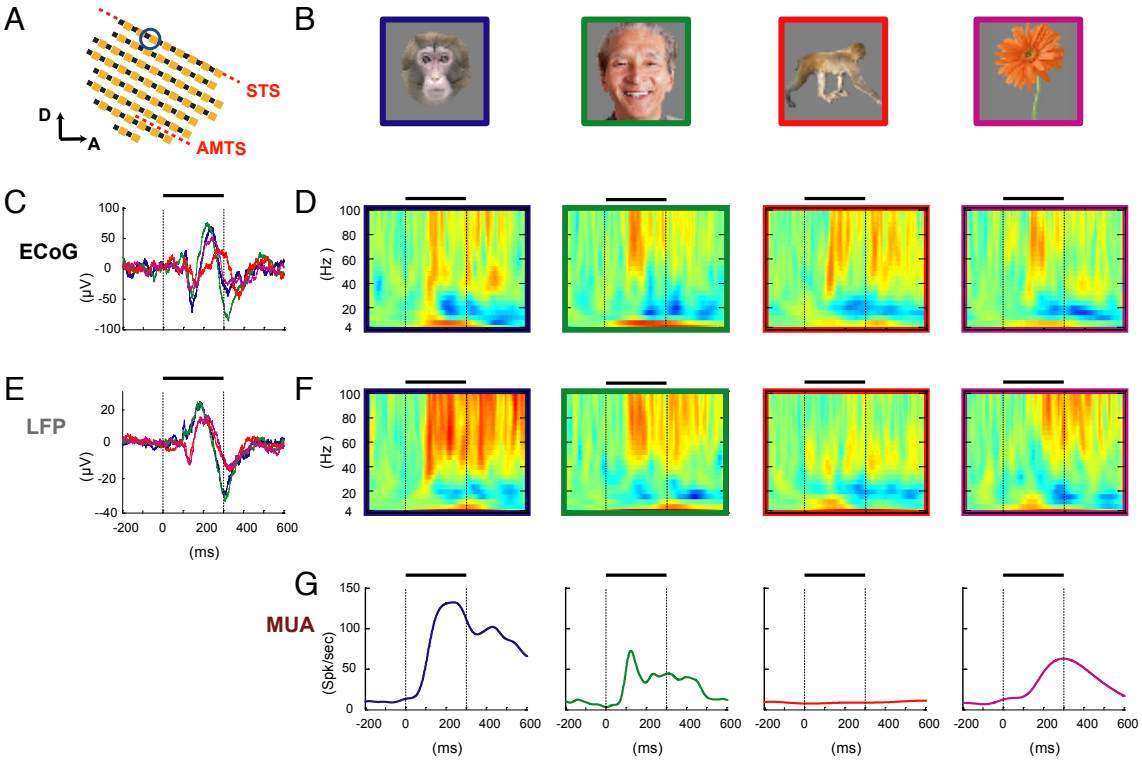

Figure 3.2. Representative visual responses of ECoG, LFP, and MUA from the same recording site. ***A,*** Position of a representative electrode. ECoG, LFP, and MUA signals were recorded from the site depicted by the blue circle, which includes a penetrating microelectrode and an adjacent ECoG electrode. ***B,*** Example stimulus images. Colored frames do not appear as part of the stimulus, but represent correspondence to the response waveforms in *C*, *E*, *G*, and the response spectrograms (see frame colors) of *D* and *F*. ***C*** and ***E*** are ECoG and LFP response waveforms. Stimuli were presented in the period between the dotted lines. Horizontal bars also show the stimulus presentation period. ***D*** and ***F*** are power spectrum responses of ECoG and LFP signals. ***G,*** MUA responses.

3.2 Materials and Methods

3.2.1 Experimental data

Data were recorded in Niigata University School of Medicine under the supervision of Dr. Miyakawa, Dr. Sawahata, Dr. Kawasaki, Dr. Matsuo, and Dr. Hasegawa. Their recording conditions were briefly described in the following sections (see Miyakawa et al. (2011) for the experimental details).

Recording electrodes and the implant surgery
For ECoG recordings, an ECoG electrode mesh (Fig. 3.1C) was designed to contain 64-channel recording electrodes plus 2 reference and 2 ground electrodes. The inter-electrode distance was 1.2 mm. The whole mesh size was 6 $\times$ 6 mm. Details on the ECoG manufacturing process have been described in previous publications (Takeuchi et al., 2005; Toda et al., 2011).Sharp microelectrodes for MUA and LFP recordings penetrated through the ECoG sheet (Fig. 3.1A). ECoG contacts and microelectrodes were arranged in the same spacing and configuration, but shifted by half the spacing distance. Electrodes were placed on area TE of the IT cortex, covering the IT gyrus and extending slightly below AMTS (Fig. 3.1A inset, 1B).

Animal task
Two Japanese macaque monkeys (Macaca fuscata), one male (9.5 kg) and one female (5.7 kg), were trained in a visual fixation task (Fig. 3.3A) to keep their gaze within $\pm 1.5^\circ$ visual angle of a fixation target (0.2–0.3° diameter). The fixation target was displayed on a 22 inch CRT monitor (Mitsubishi Electric, Tokyo, Japan) at a viewing distance of 57 cm. After 300 ms of stable fixation, a stimulus image was presented for 300 ms followed by a 600–900 ms blank interval. Two or three stimuli were presented successively in a single fixation session. Monkeys were rewarded with a drop of apple juice for maintaining fixation over the entire duration of the trial. Eye movements were captured with an infrared camera system at a sampling rate of 60 Hz.

Visual stimuli

The visual stimuli consisted of images belonging to one of 3 discrete "coarse" categories (face, body, and inanimate object) (Fig. 3.3*B* left, *C*). For the classification (decoding) analysis, images belonging to the face category were further divided into subcategories ("fine" categories) in each of three ways. For the first analysis with subcategories, the face images were divided in terms of species. The stimulus set included faces of monkeys and faces of humans (Fig. 3.3*B* right). They were labeled based on their species and the labels were predicted by neural decoders on a single-trial basis ("species classification") (see *Decoding procedure*). For the second analysis, the human face images were further divided in terms of the angles of the faces. The stimulus set included face images with three different angles (Fig. 3.3*B* right). They were labeled based on their angles and those labels were predicted by decoders ("view classification"). For the third analysis, the same human face images were divided in terms of their identity. The stimulus set included face images of five humans (Fig. 3.3*B* right). They were labeled based on their identity and those labels were predicted by decoders ("identity classification").

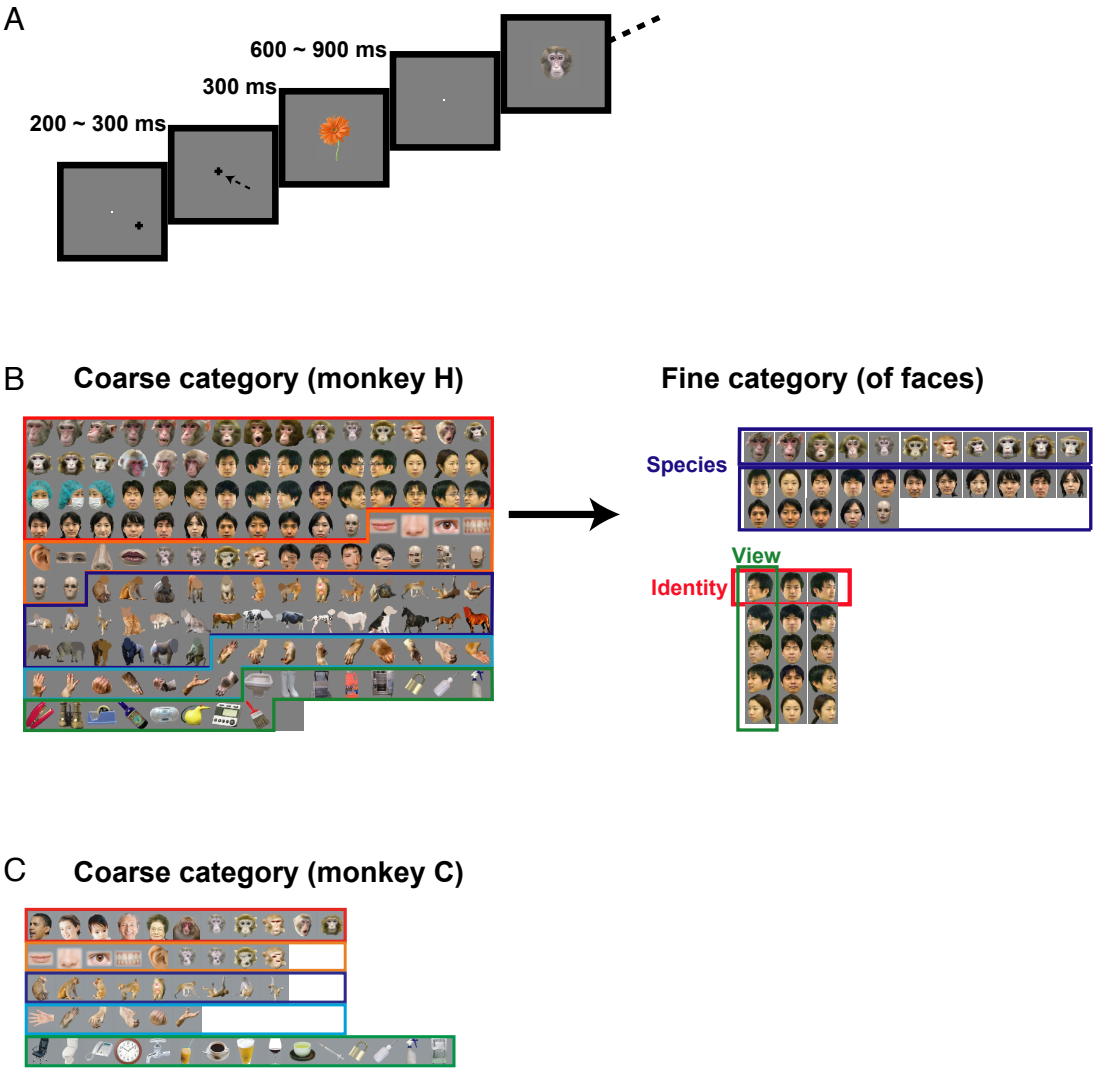

Figure 3.3. Experimental design and visual stimuli. ***A,*** Time course of presentation. Visual stimuli were sequentially presented to each monkey. Each presentation was 300 ms, followed by a 600–900-ms interval. ***B,*** Visual stimuli for monkey H. Object images including faces, body parts, and inanimate objects were used as visual stimuli for monkey H. Within face images, faces of monkeys and humans were included. Human face images could be further divided into finer subcategories in terms of facial view and identity. ***C,*** Visual stimuli for monkey C. Object images including faces, body parts and inanimate objects were used as visual stimuli for monkey C.

3.2.2 Signal features for decoding analysis

Features from MUA signals
The frequencies of spiking activity (firing rates) of the MUA were used as input features for classification. Features were extracted from MUA signals during a period from -50 ms to 600 ms relative to the stimulus onset in each trial. The signal at each microelectrode was sampled by a 100-ms time window shifted by 50 ms, and the spike rate in each time window was calculated. The spike rates of all microelectrodes and the twelve consecutive time windows were used as the input features to a decoder.

Features from ECoG and LFP signals
For ECoG and LFP data, spectral powers of the signals were used as input features. Features were extracted from ECoG/LFP signals during a period from -50 ms to 600 ms relative to the stimulus onset in each trial. The amplitudes at each channel were sampled by a 100-ms time window shifted by 50 ms, and the spectral powers of 5 frequency bands (1–10, 10–20, 20–30, 30–60, 60–100 Hz) in each time window were calculated. To do a fair comparison with MUA, the ECoG (LFP) electrodes that overlayed (matched) microelectrodes that could not obtain good MUA signals were excluded. The powers for the five frequency bands, the consecutive twelve time windows and all remaining elecctrodes were used as input features for classification analysis.

3.2.3 Decoding procedure

A neural decoding approach was used to determine the extent to which visual object information was extracted from single trial signals for ECoG, LFP, and MUA. Decoding performance for each signal modality was evaluated by pairwise decoding analysis. A pair of object categories was selected and trials including images from the selected categories were used to train a binary classifier (decoder) to predict the category of a presented image on a trial-by-trial basis (cf., Kamitani and Tong, 2005). This procedure was applied to all pairs of the three coarse categories, all pairs of the three view categories, all pairs of the five identity categories, and the pair of the two species categories. Each binary decoder consisted of a linear support vector machine (Vapnik, 1998) (implemented by LIBSVM;

Chan and Lin, 2011).

Before decoder training, a feature normalization procedure and a feature selection procedure were applied. In the feature normalization procedure, the values of each feature were z-transformed using the sample mean and standard deviation calculated with the training dataset. In the feature selection procedure, the dimensionality of the feature vector was reduced by selecting informative features based on a univariate analysis (F-statistics) applied to the training dataset. The features were ranked by the F-value that indicated differential responses to the categories, and the top 100 features were used as inputs to the decoder. In case that the number of original features for classification was equal to or less than 100, this feature selection procedure was omitted and all features were used.

Decoding performance was evaluated by a cross-validation analysis. To evaluate generalization performance for category classification across different exemplars, we ensured that trials corresponding to the same visual stimuli were not included in both the training and test data sets (Vindiola and Wolmetz, 2011). For each category pair, N exemplars were randomly chosen per category. N was set to the number of the exemplars of the category that had fewer exemplars than the paired category. The N times 2 exemplars were divided into N groups, each of which contained two exemplars from the two different categories, then the corresponding trials were divided into N groups. $(N-1)$ groups were then used to train a decoder and the remaining group was used for evaluating the trained decoder. This procedure was repeated until the trials from all N groups were tested (N-fold cross-validation), and the percentage of correct classification was calculated.

3.3 Results

By utilizing the MEMS technology developed in previous research (Takeuchi et al., 2005; Toda et al., 2011), Miyakawa et al. (2011) placed 64 ECoG electrodes and 64 microelectrodes on the monkey IT cortex (Fig. 3.1). In their experiments, images including faces, body parts and inanimate objects were presented to the monkeys and they simultaneously recorded ECoG, LFP, and MUA responses (Fig. 3.2,3).

In order to quantify the extractable information obtained from each recording modality, we examined whether the presented stimulus categories could be reliably predicted from single-trial ECoG, LFP, or MUA signals by a statistical classifier. Binary classification performance was calculated separately for all coarse category pairs (face VS body, body VS inanimate object, and inanimate object VS face). Performances were averaged across two animals and all pairs (Fig. 3.4*A*). The averaged coarse category classification performance was significantly above the chance level of 50% ($p < 0.0001$, t-test), with all the three recording modalities. Specifically, single-trial ECoG signals carried information for predicting the coarse category of the visual stimulus with a correct classification performance of 89.2%, which was significantly higher ($p < 0.001$, chi-square test, corrected for multiple comparison) than the performance obtained with MUA responses (86.8%), but significantly lower ($p < 0.0001$) than the performance with LFP responses (91.8%).

In addition to coarse category classifications, performance was also calculated for species category classifications, view category classifications, and identity category classifications (Fig 3.4*B*). ECoG signals contained sufficient information for predicting fine categories of stimuli. The accuracies for the view angles of faces were 73.0% (ECoG), 75.5% (LFP), and 79.2% (MUA), respectively (Fig. 3.4*B*, left). Likewise, the accuracies for the species of faces (macaque face VS human face) were 69.9% (ECoG), 78.3% (LFP), and 80.2% (MUA), respectively. ECoG signals did not carry sufficient information to predict facial identity (51.5%), while LFP (64.3%) and MUA (69.2%) signals did contain significant information for identity classification.

Overall, ECoG demonstrated better performance than MUA in coarse category decoding, but became inferior to MUA in decoding finer category information such as species, view, and identity information. LFP was the best recording modality for coarse category classification, and was superior to ECoG, but inferior to MUA in all fine category classifications.

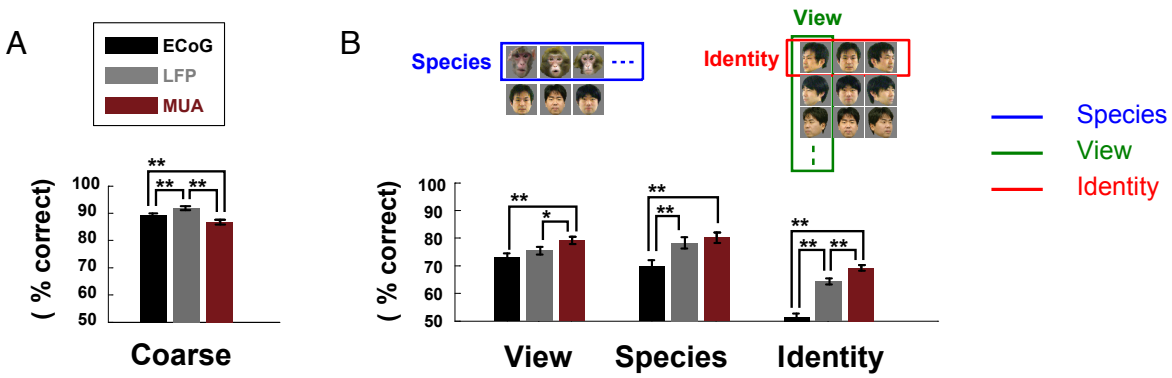

Figure 3.4. Comparison of decoding performances between ECoG, LFP, and MUA. ***A,*** Decoding performances for coarse category classification. A binary classifier was trained and tested for each pair of the three coarse image categories (face, body, and inanimate object). The averaged accuracies for the three recording methods are shown (chance level: 50%). Black asterisks show significant differences between accuracies (*: $p < 0.05$, **: $p < 0.01$; chi-square test). ***B,*** Decoding performances for fine category classification. The averaged accuracies for facial view, species and identity classifications are shown. The format is the same as in *A*.

3.4 Discussion

Miyakawa and coworkers developed a new mesh-shaped ECoG electrode array for simultaneous measurement of the three electrophysiological signals, ECoG, LFP, and MUA, under a fair condition where the numbers of electrodes and coverage for the three recording modalities match. Using their data, this thesis reports the first direct comparison of decoding performance between these recording modalities. As a result, all modalities showed significantly above-chance performances in coarse category classification (Fig. 3.4).

We found that the performance of the recording modality depended on the coarseness of the target categories of the classification. Namely, ECoG yielded better performance than MUA in coarse category decoding, but became inferior to MUA in classifying facial images for finer categories such as species, views, and identities. Overall, LFP was the best recording modality for coarse and fine classifications, except for the identity classification.

These results suggest that, although LFP provided better decoding performance than ECoG in all conditions in our analysis, ECoG signals have a significant amount of information on at least coarse categories and may be useful for revealing neural representations.

The fact that ECoG was better than MUA in coarse category classification, while MUA was better than ECoG in fine category classification, may reflect the cortical organization of category selective columns and the spatial resolution provided by each recording modality. To refer to this point further, extra analyses were conducted and summarized in our previous reports (Miyakawa et al., 2011, Miyakawa et al., in preparation).

4. Application to the human cortex: Characterization of the object-selective area

4.1 Introduction

Next, by analyzing human ECoG responses, we investigated how brain regions that respond to images with a specific object category are distributed within the ventral occipitotemporal cortex (vOT), which is considered a challenging task with fMRI. Previous fMRI studies have shown that there exist brain regions that are selectively activated by face images (Kanwisher et al., 1997; Kanwisher and Yovel, 2006) and written words (Cohen et al., 2003; Dehaene and Cohen, 2011; Price 2012), respectively. How those category selective regions are distributed in the vOT remains unsettled.

Regarding the distribution of face- and letterstring-selective regions with the vOT, three hypotheses have been proposed (Fig. 4.1). First, both face-selective and letterstring-selective vOT regions might form distinct and contiguous areas with no macroscopic spatial overlap between them (single-zone hypothesis). Second, multiple face-selective regions and multiple letterstring-selective regions might exist and alternate each other within the vOT (multizone hypothesis). Third, face-responsive and letterstring-responsive regions might be macroscopically overlapped in the vOT (overlap hypothesis).

Although category-selective regions have been discovered and investigated using fMRI, localizing those regions with enough quality to distinguish between the above three hypotheses is a challenging task. Previous studies have shown that fMRI responses in the vicinity of the vOT suffer from artifacts arising from multiple sources (Ojemann et al., 1997; Winawer et al. 2010).

A pioneering ECoG study reported that face-selective and letterstring-selective recording electrodes occasionally abutted each other in individual hemispheres (Nobre et al. 1994), which supports the multizone hypothesis. Also, other ECoG studies have found evidence for the existence of multiple face-selective regions

(Allison et al. 1999; Parvizi et al. 2012), but this is less common due to coverage issues and electrode organization (e.g. strips vs. grids).

To address these issues, Matsuo et al. (2013) implanted high-density ECoG electrodes in human patients with epilepsy, and applied a standard electrode-wise analysis to the ECoG responses to localize regions selective to faces and written words (Fig. 4.2). As a result, they found that there exist separate multiple face-selective and written word-selective regions in the human vOT (Fig. 4.3).

However, electrode-wise analysis alone may fail to quantify information encoded in ECoG response patterns over multiple electrodes. Previous studies have shown that even within a cortical region that maximally responds to one category (e.g. face), activity patterns over voxels can encode information to discriminate among other categories (Haxby et al., 2001). For this reason, it is possible that the category-selective electrodes defined by Matsuo et al. (2013) contain information for discriminating among other categories, thus their definition of category-selective electrodes may be invalid.

Decoding analysis was applied to examine whether the face-/letterstring-selective electrodes defined by Matsuo et al. (2013) have information only for discriminating the face/letterstring category, rather than other categories as a population. The face- and letterstring-selective electrodes were used as decoder inputs, and the binary classification accuracies for each category pair were calculated. We confirm that the face-/letterstring-selective electrodes show higher classification accuracies in discriminating between faces and letterstrings from other categories, compared to discriminating among the other categories.

This work has been published as part of Matsuo et al. (2013), and the author of this thesis conducted the decoding analysis in it.

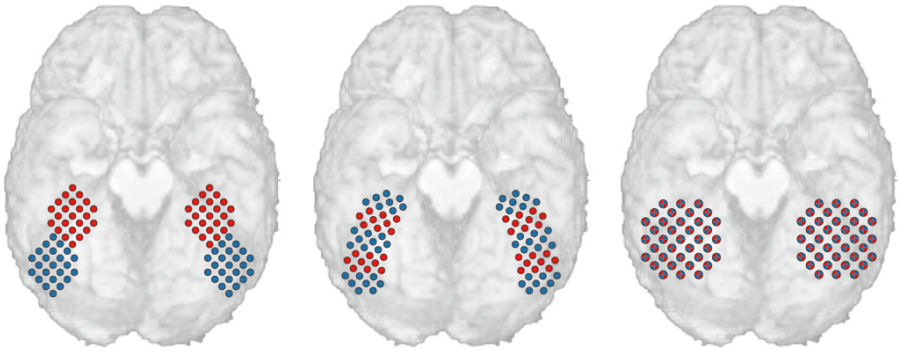

Figure 4.1. Three models of the spatial distribution of face- and letterstring-selective regions in the ventral occipitotemporal cortex (vOT). Single-zone (left), multizone (center), and overlap (right) models for face-selective (red), letterstring-selective (blue), and both face- and letterstring-selective channels are superimposed on the ventral view of a spatially normalized brain.

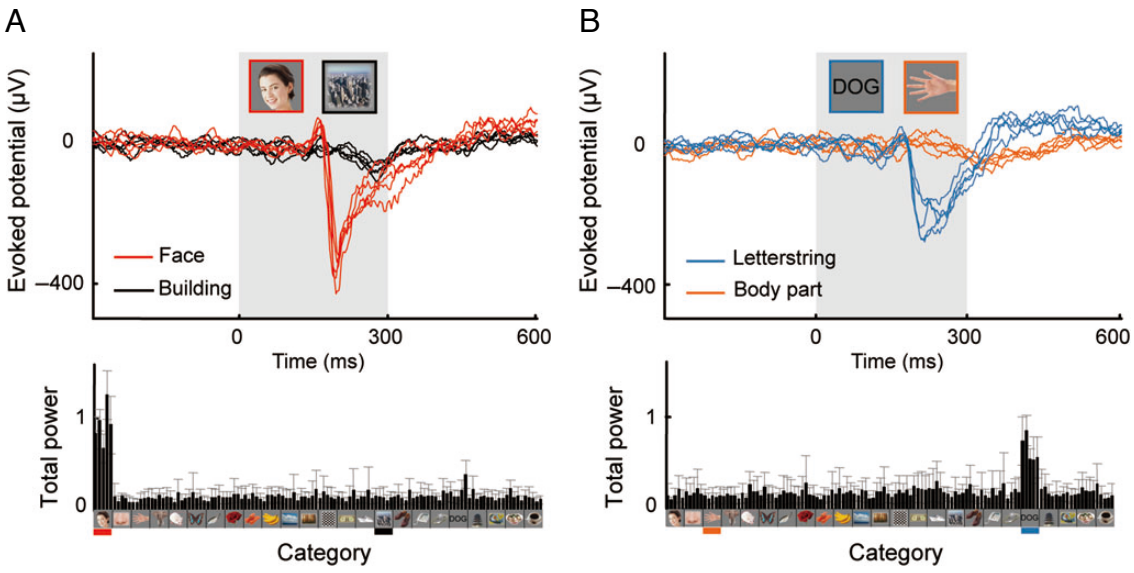

Figure 4.2. Category-selective responses for face and letterstring stimuli. ***A,*** Visually evoked potentials (VEPs) and the total electrocorticography (ECoG) power selective to the face category from a representative electrode. Upper, traces represent VEPs from different exemplars of face (red) and building (black) categories. Shadows denote stimulus presentation periods. Lower, the total ECoG power for the electrode is plotted against the stimulus category. The 5 successive bars represent evoked ECoG powers from 5 exemplar face images. Error bars indicate standard deviations (SDs). Color bars under the images denote category labels as in VEPs. ***B,*** Selectivity of a representative electrode selective to the letterstring category. Formats are the same as in *A*.

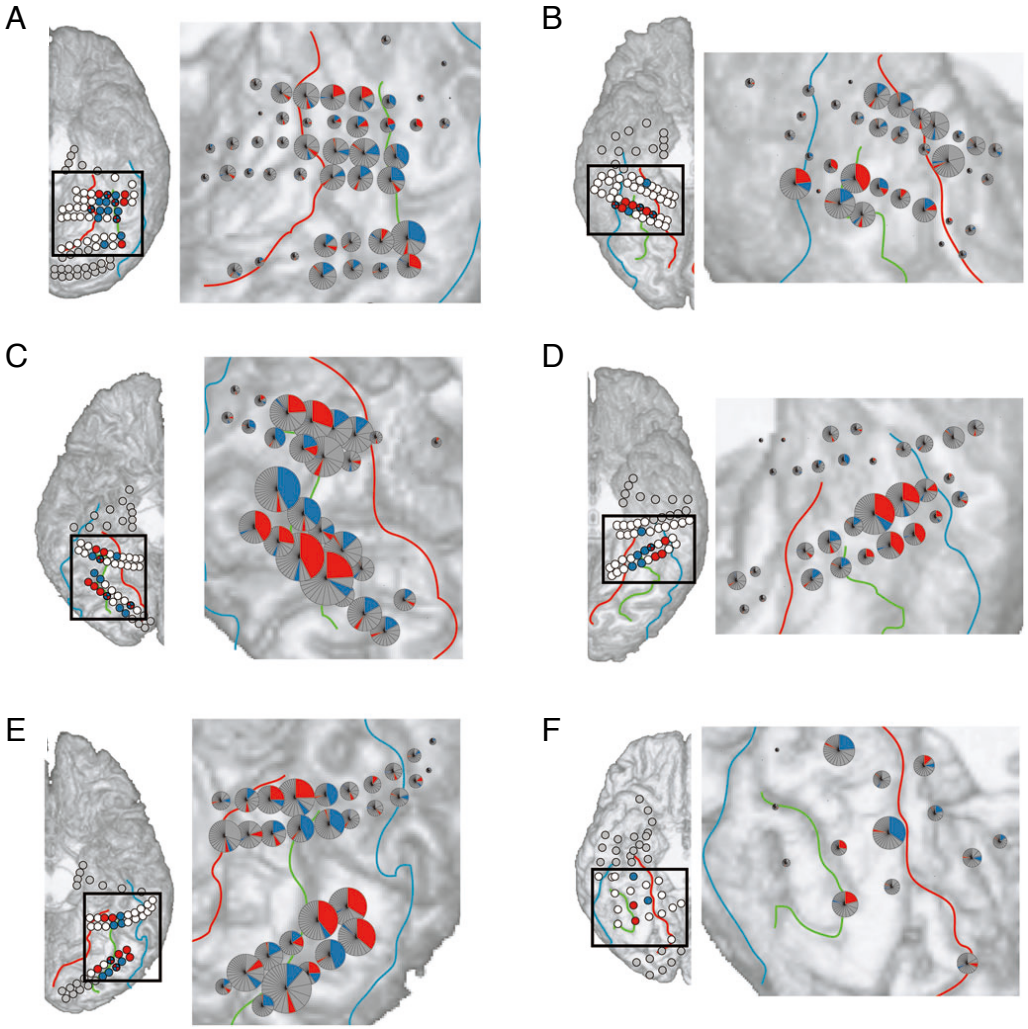

Figure 4.3. Highly face- and letterstring-selective ECoG channels in each individual hemisphere. ***A***–***F***, Left, channels significantly selective to faces (red circles), letterstrings (blue circles), and both (quarter red/blue circles) in the vOT ($-80 < y < -30$ and $z < -10$), shown on ventral views of the brain in 6 hemispheres from 4 patients. White circles indicate channels nonselective for these 2 categories and gray circles represent channels outside the vOT. Blue, green, and red lines indicate occipitotemporal, midfusiform, and collateral sulci. Right, each pie chart depicts the relative F-value, evaluated by ANOVA, for each category of stimuli during a time window of 50–300 ms after stimulus onset. Diameter of each channel reflects the total power of the channel.

4.2 Materials and Methods

4.2.1 Experimental data

Data were recorded in The University of Tokyo Hospital and Nishi-Niigata Chuo National Hospital under the supervision of Dr. Matsuo, Dr. Kawasaki, Dr. Kawai, Dr. Masuda, Dr. Murakami, Dr. Kunii, Dr. Kameyama, Dr. Saito, and Dr. Hasegawa. Their recording conditions are briefly described in the following sections (see Matsuo et al. (2013) for experimental details).

Subjects
Written informed consent was obtained from 20 patients with pharmacologically intractable epilepsy, who were evaluated for possible surgical treatment at The University of Tokyo Hospital and Nishi-Niigata Chuo National Hospital. Experimental protocols were approved by the institutional review boards of both hospitals and the Niigata University School of Medicine.

Recordings
From 2009 to 2012, Matsuo and his coworkers implanted subdural electrodes in 20 patients, totaling 30 hemispheres for the purpose of detecting epileptic foci. Among them, 22 hemispheres were excluded by their preliminary analysis because those hemispheres were not suitable to examine whether the face-selective and letterstring-selective channels were spatially separated, alternated, or overlapped (see Matsuo et al., 2013 for the details). Finally, 6 hemispheres (3 left and 3 right hemispheres) from 4 patients (15–48 years old, 3 males, 4 right-handed) fulfilled the criteria (see Matsuo et al. (2013) for details).

Subdural ECoG electrode arrays were arranged in grids or strips (Unique Medical Co., Tokyo, Japan). Each grid/strip contained 4–20 electrodes. Electrode contact was 1.5 mm in diameter with a 5-mm separation, or 3 mm in diameter with a 10-mm separation. The number and location of the recording sites in the temporal, occipital, and frontal lobes were determined exclusively by clinical criteria.

The recorded signal was amplified using a reference placed on the scalp, filtered between 0.55 and 150 Hz and sampled at 400 Hz (Nicoletone, Care Fusion, San Diego, CA, USA), or amplified using an averaged intracranial reference placed outside the epileptic focus and filtered between 0.05 and 300 Hz, and sampled at 1 kHz (EEG1200, Nihon Koden, Tokyo, Japan). All data were acquired during periods without epileptic seizure events.

Stimulus presentation
Subjects performed a one-back task or a category judgment task while colored photographs of 120 objects from 24 different categories, including human faces and letterstrings, were presented in a pseudorandom order on a 27-inch LCD monitor at a viewing distance of 57 cm. During the one-back task, the same stimuli were repeatedly presented on some trials. Data from the second presentation trials were excluded so as not to underestimate the visual response due to repetition suppression. Each stimulus, subtending 6° of the visual angle, was presented for 300 ms, followed by a 900-ms interval period.

Electrode localization
Three-dimensional (3D) T1-weighted magnetic resonance images (MRIs) of each subject's brain were obtained preoperatively. MRIs were automatically registered to postoperatively scanned computed tomography to determine electrode positions based on a normalized mutual information method using AVIZO (Visualization Science Group, Bordeaux, France). The 3D brain surface was then reconstructed using Real INTAGE (Cybernet Systems, Ltd., Tokyo, Japan). For joint presentation, the 3D brain image mounted with electrode locations was normalized to Montreal Neurological Institute coordinates via a linear scale adjustment using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/, last accessed on 11/12/13) and then transformed into Talairach coordinates.

4.2.2 Electrode-wise analysis

The analysis conducted in Matsuo et al. (2013) to define face/letterstring-selective electrodes is summarized in the following section (see Matsuo et al., 2013 for the details).

Event-related spectral perturbation (ERSP)
Using the ECoG reponse during a period from 50 ms to 300 ms relative to the stimulus onset in each trial and electrode, the spectral power was calculated for each of the θ, α, β, γ, and high-γ frequency bands (4, 8–12, 16–28, 32–64, and 68–100 Hz, respectively). A 250 ms Hamming window was applied to single trial responses before the Fourier transform. According to the convention, those power values were called event-related spectral perturbations (ERSPs).

Category selectivity
To judge whether an electrode was selective to a specific category or not, an ANOVA was applied to the ERSPs. For each frequency band and electrode, an ANOVA was conducted to compare the difference in the mean ERSP between one category and the other categories, accompanied by a Bonferroni correction. The α value for each comparison was set at 0.05 divided by the number of electrodes, frequency bands, and categories. For a given category, when an electrode showed a higher mean ERSP than the other categories at either of the five frequency bands, the electrode was judged selective to the category.

4.2.3 Decoding analysis

Using a neural decoding approach, category information encoded in a group of electrodes (face-selective electrodes or letterstring-selective electrodes) was evaluated. We selected 2 object categories and used the trials in which images included in those categories were presented to train a binary classifier to predict the category of a presented image on a trial-by-trial basis (cf., Kamitani and Tong 2005). We applied this procedure to all pairs of the 24 object categories. Each binary classifier consisted of a linear support vector machine (Vapnik 1998) (implemented by LIBSVM (Chang and Lin 2011)).

As input features to the classifiers, the ERSPs were calculated with a time window between 50 and 300 ms after the stimulus onset. Then, the ERSPs for the electrodes of a group and the 5 frequency bands were used. To avoid arithmetic overflow and underflow, the values of each feature were z-transformed using the

sample mean and SD calculated from the training dataset.

To evaluate generalization performance for category classification across different exemplars, we ensured that trials corresponding to the same visual stimuli were not included in both the training and the test datasets (similar to Vindiola and Wolmetz 2011). For each category pair, we divided the exemplars included in either category of the pair into 5 groups, each of which contained 2 exemplars from the 2 different categories, and divided the corresponding trials into 5 groups. Four groups were then used to train a decoder and the remaining group was used for evaluating the trained classifier. This procedure was repeated until the trials from all 5 groups were tested (5-fold cross-validation), and the percentage of correct classifications was calculated.

4.3 Results

Collecting data from patients who underwent high-density implantation of as many as 18–46 intracranial electrodes in vOT per hemisphere, Matsuo et al. (2013) succeeded in recording ECoG responses with enough electrode-density to distinguish between the three hypotheses about the distribution of face- and letterstring-selective regions (Fig. 4.1). By applying the conventional electrode-wise analysis to their data, several face-selective electrodes and letterstring-selective electrodes were observed in the vOT (Fig. 4.2). They plotted those electrodes on each hemisphere and their results suggest that face-selective and letterstring-selective electrodes form multiple, separate clusters with no overlap (Fig. 4.3), which supports the multizone hypothesis.

To evaluate the amount of information encoded in face-selective and letterstring-selective electrodes defined by them, neural decoding analysis was performed (see ***Materials and Methods***). For each category pair, a binary classifier was trained to predict the presented categories based on ECoG responses from the face- and letterstring-selective electrodes and the accuracy was evaluated.

Figure 4.4 shows those binary classification accuracies for each selective region. The face-selective electrodes showed a much higher performance for discriminating faces from other categories (92±9.5%, median ± IQR) than for discriminating letterstrings from other categories ($60 \pm 21\%$) ($p < 9.6 \times 10^{-38}$, Mann–Whitney U-test with Bonferroni correction) and for discriminating among the other categories ($57 \pm 14\%$) ($p < 6.1 \times 10^{-74}$) (Fig. 4.4*A* left, *B* left). Similarly, based on ECoG responses from the letterstring-selective electrodes, the performance for discriminating letterstrings from other categories (89±13%) was much higher than that for discriminating faces from other categories ($61 \pm 15\%$) ($p < 1.4 \times 10^{-38}$) and that for discriminating among the non-face and non-letterstring categories ($61 \pm 17\%$) ($p < 3.5 \times 10^{-66}$) (Fig. 4.4*A* right, *B* right).

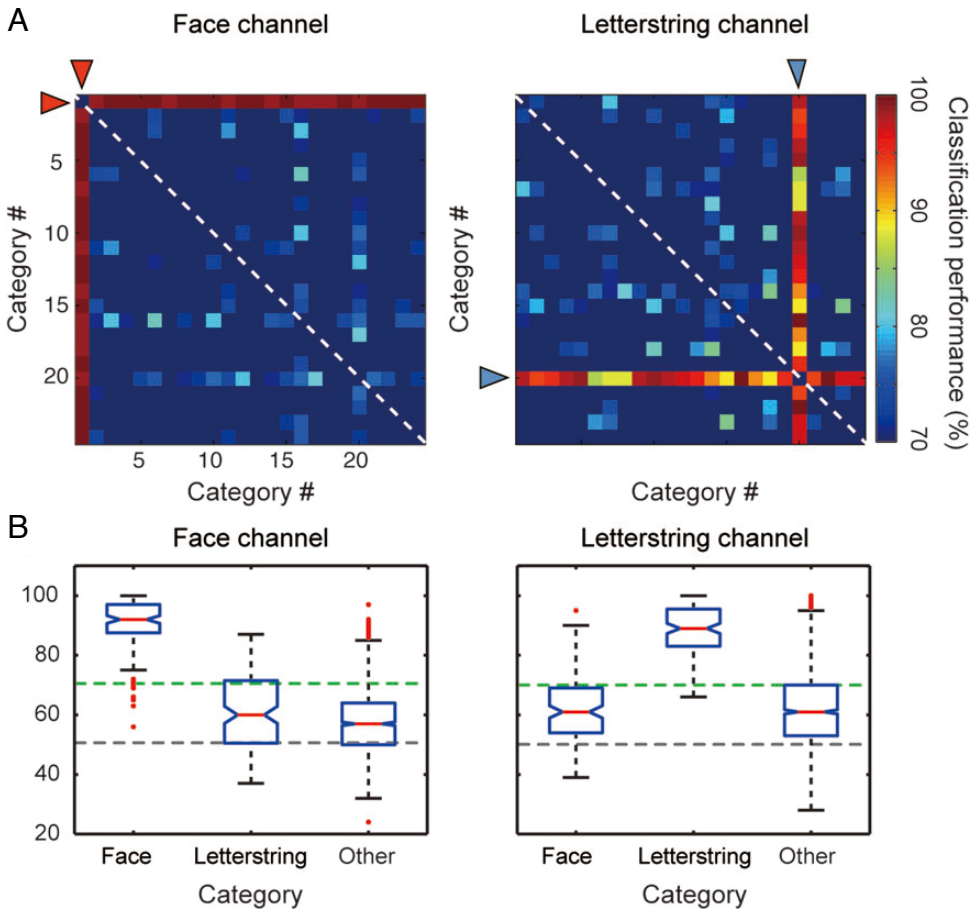

Figure 4.4. Binary classification performance based on ECoG responses from face- and letterstring-selective regions. ***A,*** Binary classification accuracies for all category pairs. For a given category pair, the decoding accuracy in discriminating between the two categories was calculated based on ECoG responses from the face-selective (left) and letterstring-selective (right) electrodes in a representative subject. Red and blue rectangles denote the face and letterstring categories, respectively. ***B,*** Box plot of averaged classification accuracies with face- and letterstring-selective electrodes over all 6 hemispheres. The accuracies in discriminating faces from the others (Face), in discriminating letterstring from the others (Lettterstring), and in discriminating among the other 22 categories (Other) are plotted. The boxplots signify the upper ($q3$) and lower ($q1$) quartiles. The median is depicted by the central red bar. The whiskers indicate the range, which include all data except outliers. Data points exceeding $q3 + 1.5 \times (q3 - q1)$ or below $q1 + 1.5 \times (q3 - q1)$ were treated as outliers and plotted individually as red dots. Broken lines indicate chance level (gray) and the threshold of significance level (green, $p = 0.01$).

4.4 Discussion

To examine how face selective- and letterstring selective-regions are distributed on the human vOT among the three hypothses (Fig. 4.1), Matsuo et al. (2013) recorded ECoG responses from the human vOT and defined face- and letterstring-responsive electrodes by a standard electrode analysis (Fig. 4.2). Their results support the multizone hypothesis that there exist separate, multiple face-responsive and letterstring-selective regions on the vOT (Fig. 4.3), which is consistent with previous ECoG studies (Allison et al. 1999; Nobre et al. 1994; Parvizi et al. 2012).

However, electrode-wise analysis alone may fail to quantify information encoded in ECoG response patterns over multiple electrodes. Previous fMRI studies have shown that even within a cortical region that maximally responds to one category, voxels can encode information to discriminate among other categories using response patterns over multiple voxels (Haxby et al., 2001). For this reason, it is possible that the category-selective electrodes defined by Matsuo et al. (2013) contain information for discriminating among other categories.

In this chapter, we used neural decoding analysis, to show that face- and letterstring-selective electrodes have more information for discriminating faces and letterstrings from the other categories than information for discriminating among the other categories as a population (Fig. 4.4).

From the electrode-wise analysis conducted by Matsuo et al. (2013) and the decoding analysis conducted here, we conclude that there exist multiple face-selective regions and letterstring-selective regions in the human vOT, and those regions are separated on the cortex.

5. Comparison of ECoG signal features: Power, Phase, and Correlation

5.1 Introduction

In many ECoG studies, spectral powers or response amplitudes in individual electrodes were often highlighted and used as input features for decoding analysis. Although a few previous studies have introduced the phases of the ECoG signals in individual electrodes (Hammer et al., 2013), efficient input signal features have not been fully explored.

According to previous studies, latency/phase differences between ECoG electrodes contain information about tasks. Theoretical studies have proposed a possibility that neural representations are formed using temporal patterns over multiple neuronal units, including the order of response latencies (Gautrais and Thorpe, 1998; Thorpe et al., 2001; Van Rullen et al., 1998), spike sequences with millisecond precision (Abeles, 1991; Abeles et al., 1993; Lestienne and Strehler, 1987; Oram et al., 1999), and gamma-band synchronization (Eckhorn et al., 1988; Engel et al., 1991; Gray et al., 1989). Experimental results have also shown that such temporal patterns encode information on low-level visual features, such as light intensity in the retina (Gollisch and Meister, 2008), orientation in the primary visual cortex (Celebrini et al., 1993; Gawne et al., 1996; Shriki et al., 2012), and co-occurrence of edges (Eckhorn et al., 1988; Engel et al., 1991; Gray et al., 1989). If such coding is adopted in a recorded region, the use of features that directly reflect latency/phase differences may improve a decoder's predictive performance.

In this chapter, using ECoG responses from human patients when they viewed objects, we examined whether the categories of viewed objects can be better predicted by using temporal correlations between electrodes. Temporal correlations measure the degree of in-phase synchronization between two time series (Varela et al., 2001). For comparison, we used the spectral power or phase values of ECoG signals in individual electrodes as features, and compare decoding performances

to examine whether temporal correlations carry additional information on object categories.

In addition, we investigated whether ECoG responses that produce informative correlations are time-locked to the stimulus or not. Informative correlations can be generated by category-specific ECoG responses with constant latencies across trials, or alternatively by relative time series of ECoG signals with variable latencies across trials. To examine this, we conducted a shuffling analysis (Averbeck et al., 2006; Yamashita et al., 2008). In this analysis, we created data in which relative timing/phase differences between electrodes within each trial were destroyed (trial-shuffled data) from the original ECoG data by randomly permuting trials with the same stimulus in each electrode. We compared the decoding performance using the shuffled data to the decoding performance of the original data.

Also, to characterize the time course of decoding performance, we calculated decoding performance as a function of time (see Liu et al., 2009) and compared this between spectral power values and temporal correlations.

Finally, to specify which encoding model explains the results of the time course analysis at the neuronal electric source level, we performed a simulation analysis in which two representative models were tested; a model where sources encode information with their response latencies (Celebrini et al., 1993; Gautrais and Thorpe, 1998; Gollisch and Meister, 2008; Shriki et al., 2012; Van Rullen et al., 1998), and a model where sources encode using their phase differences (Eckhorn et al., 1988; Engel et al., 1991; Gray et al., 1989).

From the analyses conducted here, it was shown that temporal correlations are informative features for decoding analysis compared to spectral powers and phases in individual electrodes, and this may reflect neural coding with phase differences across brain regions in the human temporal cortex.

This work has been published as part of Majima et al. (in press). The author of this thesis conducted all decoding and simulation analyses. Data were collected

in The University of Tokyo Hospital under the supervision of Dr. Matsuo, Dr. Kawasaki, Dr. Kawai, Dr. Saito, and Dr. Hasegawa.

5.2 Materials and Methods

5.2.1 Experimental data

Data were recorded in The University of Tokyo Hospital under the supervision of Dr. Matsuo, Dr. Kawasaki, Dr. Kawai, Dr. Saito, and Dr. Hasegawa. Their recording conditions are briefly described in the following sections (see Majima et al. in press for experimental details).

Subjects

Five patients with medically intractable epilepsy (5 female, 22–42 years, Table 5.1) participated in our experiments. All patients were admitted to The University of Tokyo Hospital, Japan. Patients underwent electrode implantation for the purpose of localizing seizure foci to guide neurosurgical treatment (Fig. 6.1). The locations of electrodes were determined solely by therapeutic considerations. We obtained written informed consent from the patients, and all experimental protocols were approved by the institutional review board at the hospital (#1797(3)).

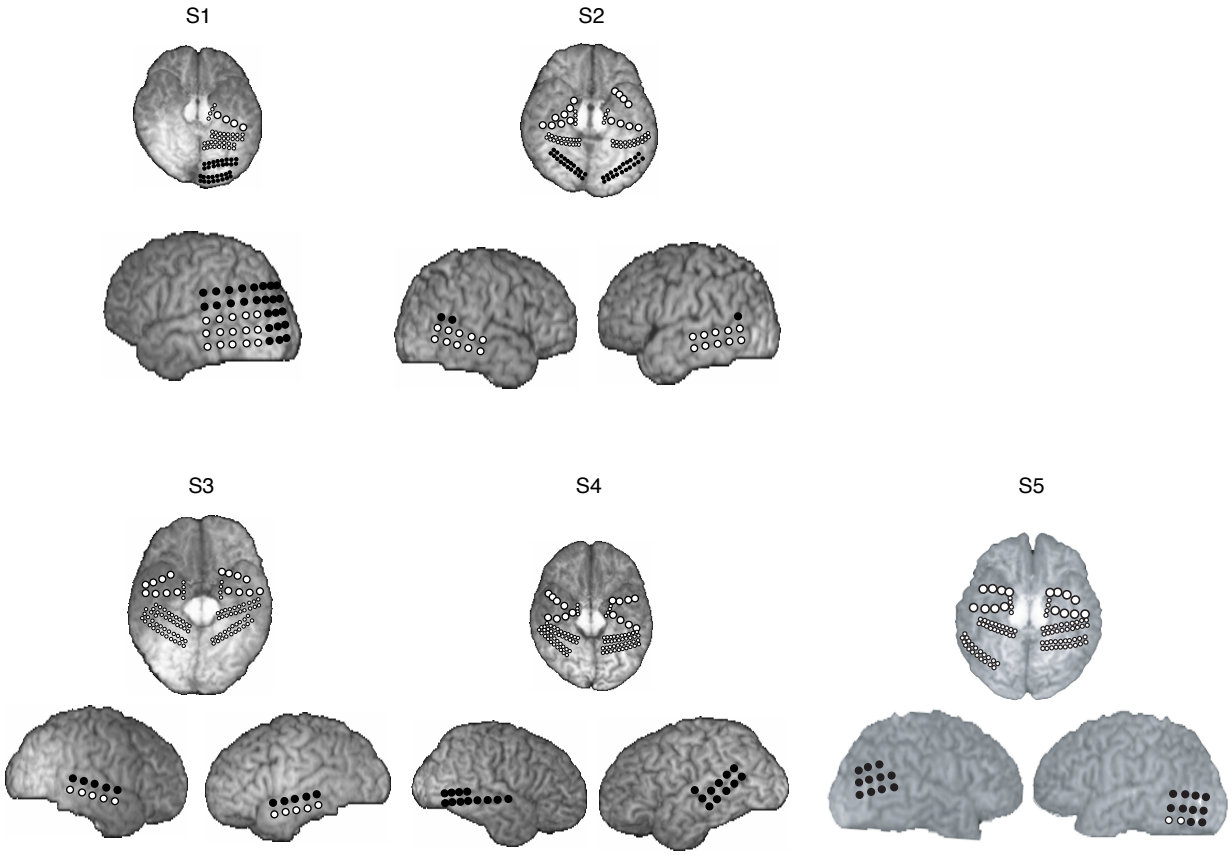

Figure 5.1. ECoG electrodes. Structural MRI images of each patient are shown with the positions of ECoG electrodes (black and white circles). Large and small circles show electrodes with contact sizes of 3.0 mm and 1.5 mm, respectively. A total of 120 electrodes for S1 and 127 electrodes for the other patients were implanted for medical diagnosis. Electrodes on the IT cortex that consists of the middle and inferior temporal gyri were selected and used for analysis (white circles).

ECoG recordings

Patients were implanted with subdural electrode arrays arranged in grids or strips (Unique Medical Co., Ltd., Tokyo, Japan). Each grid/strip contained 4–20 electrodes. Each electrode contact was 3 mm in diameter with 10 mm separation, or 1.5 mm in diameter with 5 mm separation. The number and location of the recording sites in the temporal, occipital, and frontal lobes were determined exclusively by clinical criteria. Data were obtained from a total of 628 recording sites, with 120–127 electrodes per subject. Recorded signals were amplified using a reference electrode placed on the scalp, filtered between 0.55 and 150 Hz and sampled at 400 Hz (Nicoletone, CareFusion, San Diego, CA, USA). In addition to this bandpass filtering, the signals were filtered to pass each of five frequency bands (1–10, 10–20, 20–30, 30–80 and 80–150 Hz) when calculating ECoG features in bands of interest. All data were acquired during periods without seizure events.

Electrode localization

To localize electrodes, the anatomical information of the brain provided by preoperative magnetic resonance imaging (MRI) and spatial information of the electrode positions provided by postoperative computer tomography (CT) were integrated. For each subject, the 3D brain surface was reconstructed and an automatic registration based on mutual information was performed using Avizo (Maxnet Co., Ltd., Tokyo, Japan). Electrodes on the IT cortex that consists of the middle and inferior temporal gyri were selected and used for the analysis of this study (55 electrodes for S1, 84 electrodes for S2, 117 electrodes for S3, 104 electrodes for S4, 106 electrodes for S5; Fig. 6.1, Table 5.1).

Table 5.1. Patient demographics, task performance, and the number of electrodes.

Subject	Age	Sex	Task performance (success/total)	# of all implanted electrodes	# of electrodes on IT cortex (left side/left bottom/ right side/right bottom)
S1	42	F	No record	120	55 (15/40/0/0)
S2	36	F	93/112	127	84 (10/32/10/32)
S3	22	F	109/113	127	117 (5/52/5/55)
S4	35	F	111/120	127	104 (0/52/0/52)
S5	38	F	53/55	127	106 (2/52/0/52)

Stimulus presentation

Matsuo and his coworkers prepared 120 colored photographs of objects from 24 different categories as stimuli (Fig. 5.2*A*). There were five different exemplar images per category. The stimuli were presented on a 27-inch LCD monitor at a viewing distance of 57 cm with a central fixation point (0.5°) (Fig. 5.2*B*). Each stimulus subtended a $6° \times 6°$ visual angle and was presented for 300 ms followed by a 900-ms interval period (Fig. 5.2*C*). The stimuli were presented in pseudorandom order and each stimulus was presented either 10 or 11 times per subject. The subjects were instructed to fix their eyes on the fixation point and to perform a one-back task, indicating whether an exemplar was repeated successively or not by pressing a button (55–120 repetition trials per subject). Trials with button presses were excluded from any analyses reported in this study.

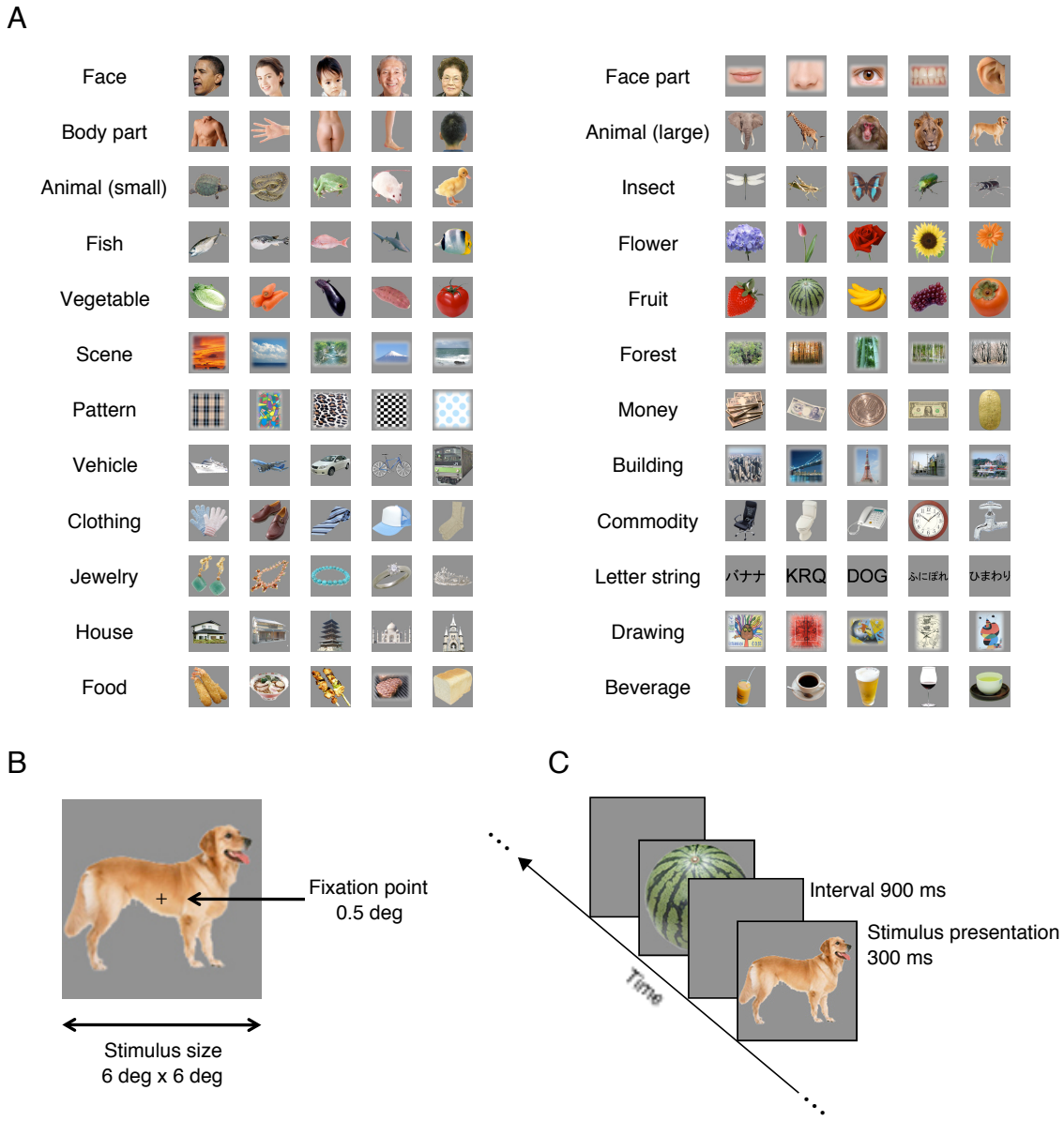

Figure 5.2. Visual stimuli and experimental design. ***A,*** Visual stimuli. We used 120 colored photographs of objects from 24 different categories as visual stimuli. There were five different exemplars per category. ***B,*** Stimulus presentation. A visual stimulus ($6° \times 6°$) was shown with a gray background and a central fixation point ($0.5°$). ***C,*** Time course of presentation. Visual stimuli were sequentially presented to subjects. Each presentation was 300 ms long followed, by a 900-ms interval.

5.2.2 Signal features for decoding analysis

Signal features used as input to decoding analyses were calculated from single-trial time series of the individual electrodes in the temporal cortex within a time window from 0 to 300 ms relative to the stimulus onset or a 100/300-ms sliding time window with a step size of 25 ms.

"Power features" were the power values of five frequency bands (1–10, 10–20, 20–30, 30–80, and 80–150 Hz) calculated from the power spectrum for each time window (the number of electrodes times the five frequency bands).

"Correlation features" were the correlation coefficients of time series for all pairs of electrodes within each time window (1485–6786 features for each subject). In the analysis to compare specific frequency bands (the five bins of power spectral frequencies), power features were limited to those in the frequency bands of interest. Correlation features were re-calculated from the band-pass filtered time series.

Correlation features are higher-order variables based on the products of (normalized) signal amplitudes between electrodes at each time point. As a result, the number of available features is much larger than that of power features. To control for the effect of higher-order feature extraction and the number of available features, the products of power values were used as an additional feature type ("product-of-power features"). In each frequency band, the products of the power values were calculated for all pairs of electrodes (five [frequency bands] times the number of correlation features).

The phase values of ECoG time series were also used as features (Hammer et al., 2013; Lopour et al., 2013). The phase values of five frequencies (6, 16, 26, 56, and 116 Hz; near-middle values of the five frequency bands) in each time window were calculated using the fast Fourier transform (FFT), and transformed to the sine and cosine values. These values were used as "phase features" (the number of electrodes $\times$ 5 [frequencies] $\times$ 2 [trigonometric functions]). Note that individual phase features do not explicitly indicate relative timings between electrodes,

although the linear decoder may detect relative timings via the weighted summation of phase features.

In the shuffling analysis to characterize informative correlations, correlation features were calculated with ECoG data in which relative timing/phase differences between electrodes within each trial were destroyed by a trial-shuffling procedure (Averbeck et al., 2006; Yamashita et al., 2008). In each electrode, the data were randomly permuted across trials for the same visual stimulus, while preserving the ECoG time series within each electrode and trial. In these shuffled data, if ECoG responses to each category were constant across trials (time-locked to the stimulus onset), the original correlations would be preserved (Fig. 5.3*A*). However, if the original correlations were from the relative time series in each trial (not time-locked to the stimulus onset) and the variability of response latencies across trials was sufficiently large, the correlations would be removed by the shuffling (Fig. 5.3*B*).

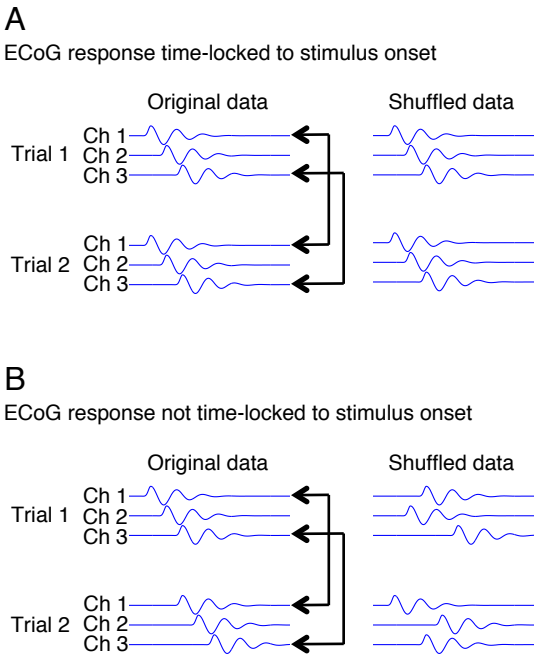

Figure 5.3. Effect of trial shuffling on ECoG responses time-locked and unlocked to the stimulus onset. ***A,*** Effect on time-locked responses. Multi-electrode ECoG responses specific to a stimulus arise with constant latencies across trials. In this case, the original response pattern remains after permuting trials in each electrode. ***B,*** Effect on unlocked responses. Multi-electrode ECoG responses specific to a stimulus arise with different latencies across trials while preserving relative timings between electrodes. In this case, the original response pattern is destroyed after permuting trials in each electrode.

5.2.3 Decoding procedure

A linear classifier (decoder) was constructed to predict the categories of presented stimuli from ECoG feature vectors on a trial-by-trial basis. Feature vectors labeled by the stimulus categories of individual trials in each subject were divided into training and test datasets, and a linear support vector machine (SVM; Vapnik, 1998) algorithm determined the parameters of the decoder using the training dataset. The decoder calculated the linearly weighted sum of the features plus a bias for each category (class) given a feature vector from the test dataset. The category with the maximum value was chosen as the predicted category (Kamitani and Tong, 2005).

Before decoder training, a feature normalization procedure and a feature selection procedure were performed. In the feature normalization procedure, the values of each feature were normalized using the sample mean and standard deviation calculated from the training dataset. In the feature selection procedure, the dimensionality of the feature vector was reduced by selecting informative features based on a univariate analysis (F-statistics) applied to the training dataset. The features were ranked by the F-value that indicated differential responses to the categories, and the top N features were used as inputs to the decoder. The number of used features (N) was decided by a cross-validation analysis within the training dataset, where N was varied from 50 to 1000 in increments of 50, and the N with the highest accuracy was chosen (nested cross-validation). The average numbers of selected features for the analysis with the 0–300-ms time window were 70 ± 63 (mean $\pm$ SD across subjects) for power features, 694 ± 404 for correlation features, and 778 ± 330 for the combined features.

To see the dependence on the number of features, cross-validation analysis with a fixed N (without the nested cross-validation to optimize N) were performed. Decoding performance was calculated for each feature type while N was varied from 10 to 1000, incremented by 10.

To evaluate generalization performance for category classification across different exemplars, it was ensured that trials corresponding to the same visual stimuli

were not included in both the training and test datasets (Vindiola and Wolmetz, 2011). The 120 stimuli were divided into five groups, each of which contained 24 stimuli from the 24 different categories. Then the corresponding trials were divided into five groups, four groups of which were used to train a decoder and the remaining group was used for evaluation. This procedure was repeated until the trials from all five groups were tested (5-fold cross-validation), and the percentage of correct classifications was calculated. The cross-validation for determining the number of features was performed using the four groups in the training dataset (4-fold cross-validation).

To compare decoding accuracy between conditions, a chi-square test and a paired t-test were used for within-subject analysis and group analysis, respectively. For the group analysis, the accuracy of each subject was logit-transformed, $\log(a/(1-a))$ where a is the classification accuracy, and then a paired t-test was applied to those values. Logit-transformed accuracies were used rather than the original ones because a normal distribution has infinite support while accuracies are bounded from zero to one, and the assumption that obtained accuracies follow a normal distribution is not suitable (Lesaffre et al., 2007). We could have used a hierarchical model to account for a binomial distribution of the accuracy in individual subjects, but since many trials were available in individual subjects (1200 trials per subject) and the standard error of the mean accuracy was very small ($1.1 \pm 0.1\%$), the measured accuracy was assumed to be the mean accuracy in each subject.

5.2.4 Simulation analysis

To perform a simulation analysis, 200 neural signal sources were arranged in a one-dimensional array with 1-mm intervals. Ten ECoG electrodes were placed 1 mm above those sources with 10 mm intervals. The signals of the ECoG electrodes were expressed by the spatial sums of the sources with a lead field matrix (Nunez, 1981).

In the latency-coding model, each source had a fixed waveform produced by sampling from an i.i.d. Gaussian distribution ($N(0, 3^2)$). Given a stimulus, each

source produced a waveform with a latency specific to the stimulus category. Those latency values were uniformly distributed from 130 to 180 ms over stimulus categories. In each trial, the latencies of all sources were jittered by the same amount randomly chosen from between 0 and 50 ms (while preserving relative timings across trials), and independent Gaussian white noise ($N(0, 3^2)$) was added.

In the phase-reset model, the sources were oscillators at a frequency of 40 Hz with an amplitude of one. In each trial, the initial phase values of all sources were randomly chosen. The phases values were then reset to values specific to the stimulus category at the reset time, which was randomly chosen for each trial between 130 and 180 ms. Independent Gaussian white noise ($N(0, 3^2)$) was added to all sources in each trial.

The number of categories was set to 24 and each category had 50 trials ($24 \times 50 = 1200$ trials in total), to match the empirical data. Electrode signals were filtered between 0.55 and 150 Hz, which was similar to the recording condition in the experiments. Using simulated data produced by each model, decoders were constructed using either spectral power or correlation features, and the decoding performance was calculated as a function of time.

In addition, modified versions of the two models to take into account category-specific amplitude modulation of sources were also tested. It is commonly assumed that the amplitudes of neuronal sources change depending on the stimulus (Naruse et al., 2010; Palva and Palva, 2007). In the latency-coding model, each source signal before noise addition was multiplied by a value specific to the given stimulus category, which was randomly chosen from one to $(M + 1)$ in each category, for each electrode. In the phase-reset model, the amplitude of each source was changed from one to a value specific to the given stimulus category at 130 ms. The amplitudes after 130 ms were multiplied by category-specific values randomly chosen from one to $(M+1)$. The parameter M was manipulated to control the degree of dominance of amplitude coding in the model. When $M = 0$, both models had no amplitude modulation.

Power and correlation features were calculated from a 100-ms sliding time window. The correlations for all electrode pairs and the power values for all electrodes and the five frequency bands were used as input features with no feature selection.

5.3 Results

Collaborators in The University of Tokyo Hospital measured ECoG data from implanted electrodes located predominantly in the temporal cortex; data from five subjects were recorded while those subjects sequentially viewed natural images (Fig. 5.2). It was confirmed that all data were acquired during periods without seizure events. Electrodes on the IT cortex that consists of the middle and inferior temporal gyri were selected and used for analysis in this study (Fig. 6.1, Table 5.1).

5.3.1 Categorical selectivity of features with representative electrodes

To illustrate the category selectivity of spectral power and temporal correlations, we calculated the trial-averaged time courses of power values in the 1–10 Hz spectral band and that of temporal correlations responding to each stimulus category. Figure 5.4 shows the time courses of the power values at two representative electrodes for the face category and the letterstring category, and those of the temporal correlation between the electrodes. In this example, the temporal correlation showed category-selective time courses whereas the modulation of power at the two electrodes is not distinctive.

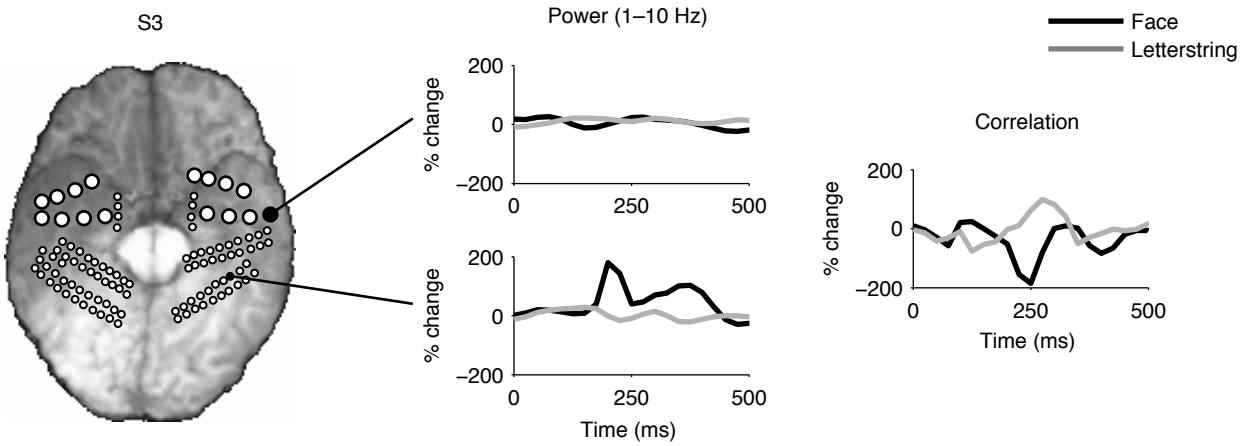

Figure 5.4. Category selectivity in power and correlation with representative electrodes. The trial-averaged time courses of the power values in the 1–10 Hz band at two electrodes (middle) and of the correlations (right) are shown for the face and letterstring categories.

5.3.2 Comparison of decoding performance across feature types

To evaluate information encoded in each type of feature, we performed decoding analysis using the power, correlation and combined features calculated from the ECoG time series from 0 to 300 ms relative to the stimulus onset (24 categories; chance level, 4.16%). Fig. 5.5*A* shows the cross-validated performances for the three feature types in each subject, using the number of features optimized via nested cross-validation (see ***Materials and Methods***). Performance exceeded the chance level for all feature types and all subjects ($p < 10^{-5}$, binomial test). By combining power and correlation features, the decoding performance was substantially improved compared with the performance using power alone ($p < 0.05$, paired t-test across subjects; $p < 0.05$ in all individual subjects, chi-square test). Even correlation features alone significantly outperformed power features ($p < 0.05$, paired t-test across subjects; $p < 0.05$ in four out of five subjects, chi-square test).

Fig. 5.5*B* shows the mean performances as a function of the number of used features. In addition to power, correlation, and combined features, the results for the products of power values (product-of-power features) are shown, which were introduced to control for the effect of higher-order feature extraction and the number of available features (see ***Materials and Methods***). The performances with power and product-of-power features peaked around 50 features, while the performances with correlation and combined features continued to improve, even with > 500 features. When compared with the same number of features used, correlation and combined features outperformed power features even at small numbers of features (around 50–250 features).

Furthermore, the decoding performance with product-of-power features remained lower than those with the other features, even though the number of available features was the largest. Thus, the performance improvement by adding correlation features (Fig. 5.5*A*) cannot be attributed to the larger number of used or available features, alone. Temporal correlations appear to carry additional information about object categories that is not represented by powers or higher-order features generated from powers.

Correlation features reflect phase or timing differences of ECoG time series between electrodes. To see whether phase information in individual electrodes is sufficient to achieve the high decoding performance obtained with correlations, decoding accuracies using phase features were analyzed (the number of features were optimized by nested cross-validation; see ***Materials and Methods***). Fig. 5.6 shows the mean performances obtained with power, phase, and correlation features. The decoding performance obtained with correlation features was better than the other features ($p < 0.05$, paired t-test). Although phase features could implicitly encode phase or timing differences between electrodes, this result suggests that a more explicit representation of the differences by correlations was critical for achieving the high decoding performance.

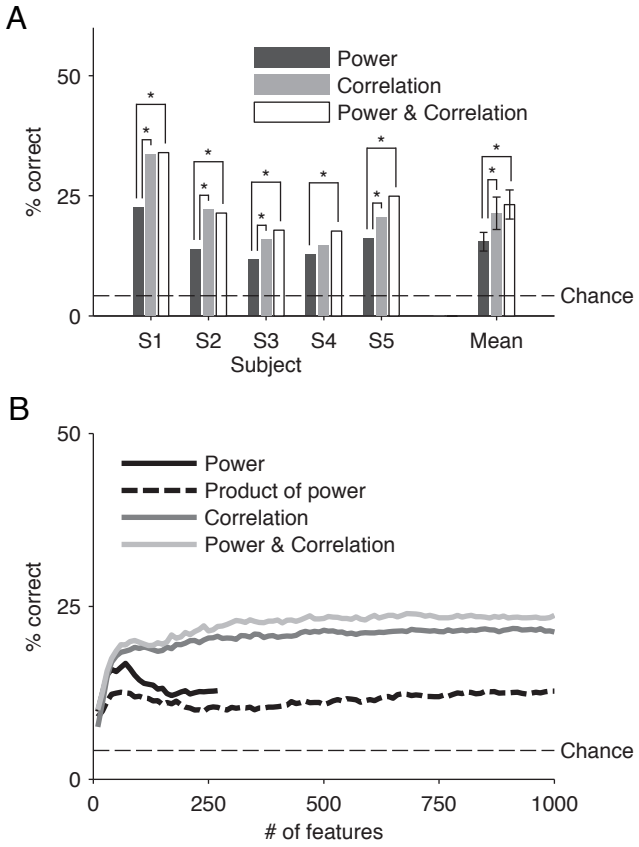

Figure 5.5. Decoding performance obtained with powers, correlations, and both. ***A,*** Results from individual subjects and the means (dashed line: chance level; *: $p < 0.05$; chi-square test on individual subjects, and paired t-test on mean performance; error bar: s.e.m. over subjects). The number of features used was determined by a nested cross-validation procedure. ***B,*** Results with different numbers of features. In addition to the above three types, the performance of product-of-power features is plotted (means across subjects). Note that the number of available features was fewer for power (< 275) than for the other feature types.

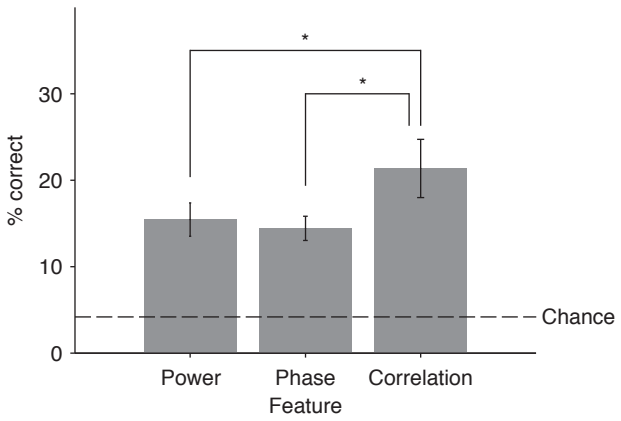

Figure 5.6. Decoding performance obtained with power, phase, and correlation features. The values for power and correlation features are the same as those in Fig. 5.5*A* (dashed line: chance level; *: $p < 0.05$, paired t-test, error bar: s.e.m. over subjects).

5.3.3 Shuffling analysis

Informative correlations could be generated by category-specific ECoG responses with constant latencies across trials (stimulus-locked responses), or alternatively by relative time series of ECoG signals with variable latencies across trials. To examine which type of response produced informative correlations, we created trial-shuffled data (see ***Materials and Methods***) and compared the decoding performance between the original and shuffled datasets.

Fig. 5.7 shows the decoding performance using the original and shuffled data for each frequency band. Using the data before the band-pass filtering ("All"), the shuffling degraded the performance in most subjects ($p < 0.05$ in 4/5 subjects, chi-square test). Among the five frequency bands, the 30–80 Hz and the 80–150 Hz bands showed performance degradation ($p < 0.05$ in 5/5 and 2/5 subjects, respectively). At the group level, the decoding performance in the 30–80 Hz was significantly degraded by shuffling ($p < 0.05$, paired t-test, corrected for multiple comparison). While the overall performance was highest in the 1–10 Hz band, the performance was comparable between the original and shuffled data. The 10–20 Hz and 20–30 Hz bands showed poor performance with both the original and shuffled data.

These results suggest that stimulus-unlocked, relative time series may contribute to category coding in the high frequency bands. Informative correlations in the low frequency band may arise in a stimulus-locked manner, but it is also possible that the variability of response latencies across trials was too small compared to the low frequency cycle to destroy correlations from relative time series.

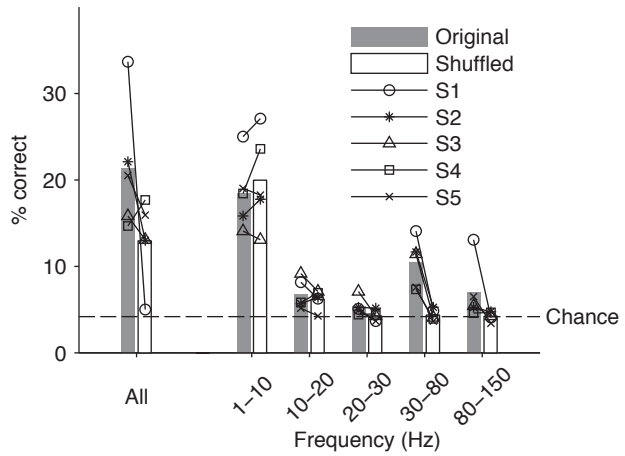

Figure 5.7. Shuffling analysis. Decoding performance obtained using correlation features is compared between the original and shuffled data. The analysis was applied to ECoG signals before band-pass filtering (all) and band-pass filtered signals. The bar graph shows the mean performance over subjects (dashed line: chance level). Symbols represent individual subjects.

5.3.4 Time course of decoding performance

To characterize the time course of decoding performance, the decoding performance was calculated as a function of time using a sliding time window. The time window was shifted by 25 ms, and decoding performance was plotted for the power and correlation features (Fig. 5.8, upper). In this analysis, power values from five frequency bands and correlations for the raw ECoG signals were used as input features, respectively. The decoding performance with the 300-ms time window began to rise when the center of the time window was within 100 ms after the stimulus onset and remained above chance level even after the disappearance of the stimulus (300 ms).

To evaluate when decoding performance began to rise from chance level, we defined the onset time as the first point at which the mean decoding performance over subjects exceeded the 99% percentile (% correct = 5.4%) in the performance distribution at 0 ms. The decoding performance with correlations rose earlier than that with powers (25 ms and 100 ms for correlation and power features, respectively).

To test the dependence on the width of the time window used, a 100-ms sliding time window was also tested (Fig. 5.8, lower). Even with the shorter time window, the onset time for correlations was earlier than that for powers. The peak performance for correlations was slightly lower than that for powers with this short time window, presumably because correlations could be better estimated using a longer time series.

The results indicate the efficacy of temporal correlations between electrodes to encode categorical information in early responses.

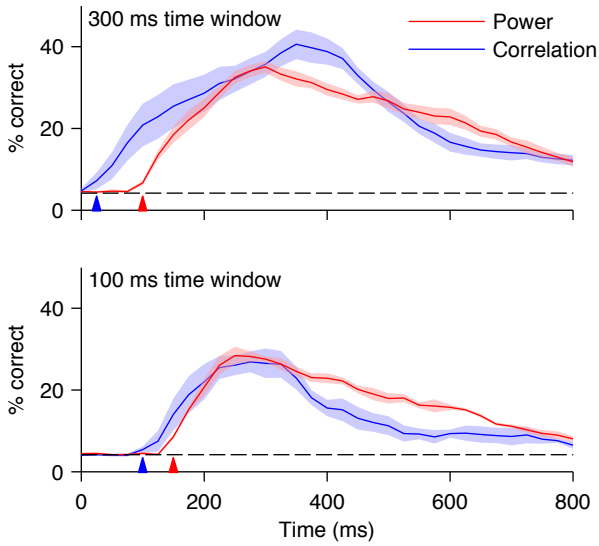

Figure 5.8. Time course of decoding performance. Decoding performances obtained with power (red) and correlation (blue) features are plotted as a function of time for 300-ms and 100-ms time windows (dashed line: chance level; shaded region: s.e.m. over subjects). Each arrow indicates the center of the time window at which the decoding performance exceeded the threshold.

5.3.5 Simulation analysis

To examine what mechanism could explain the early onset of the good decoding performance with correlations, two representative models were tested: a latency-coding model and a phase-reset model, where category-specific temporal patterns were encoded by latency differences and by phase differences across sources, respectively (Figs. 5.9*A*, *B*; see ***Materials and Methods***; Eckhorn et al., 1988; Engel et al., 1991; Gautrais and Thorpe, 1998; Gray et al., 1989; Thorpe et al., 2001; Van Rullen et al., 1998).

In the latency-coding model, decoding performance began to rise around the same time for both power and correlation features in all conditions (Fig. 5.9*C*), failing to account for the dissociation between power and correlation features. This is presumably because changes in power and correlation features are tightly coupled via response latencies.

In contrast, for the phase-reset model the performance with correlations began to rise earlier compared with power, except when the amplitude modulation was strong ($M = 10$; Fig. 5.9*D*, bottom). Phase resetting may cause only subtle differences in power via the summation of synchronized/desynchronized neuronal sources, resulting in a slower onset of decoding performance with power features.

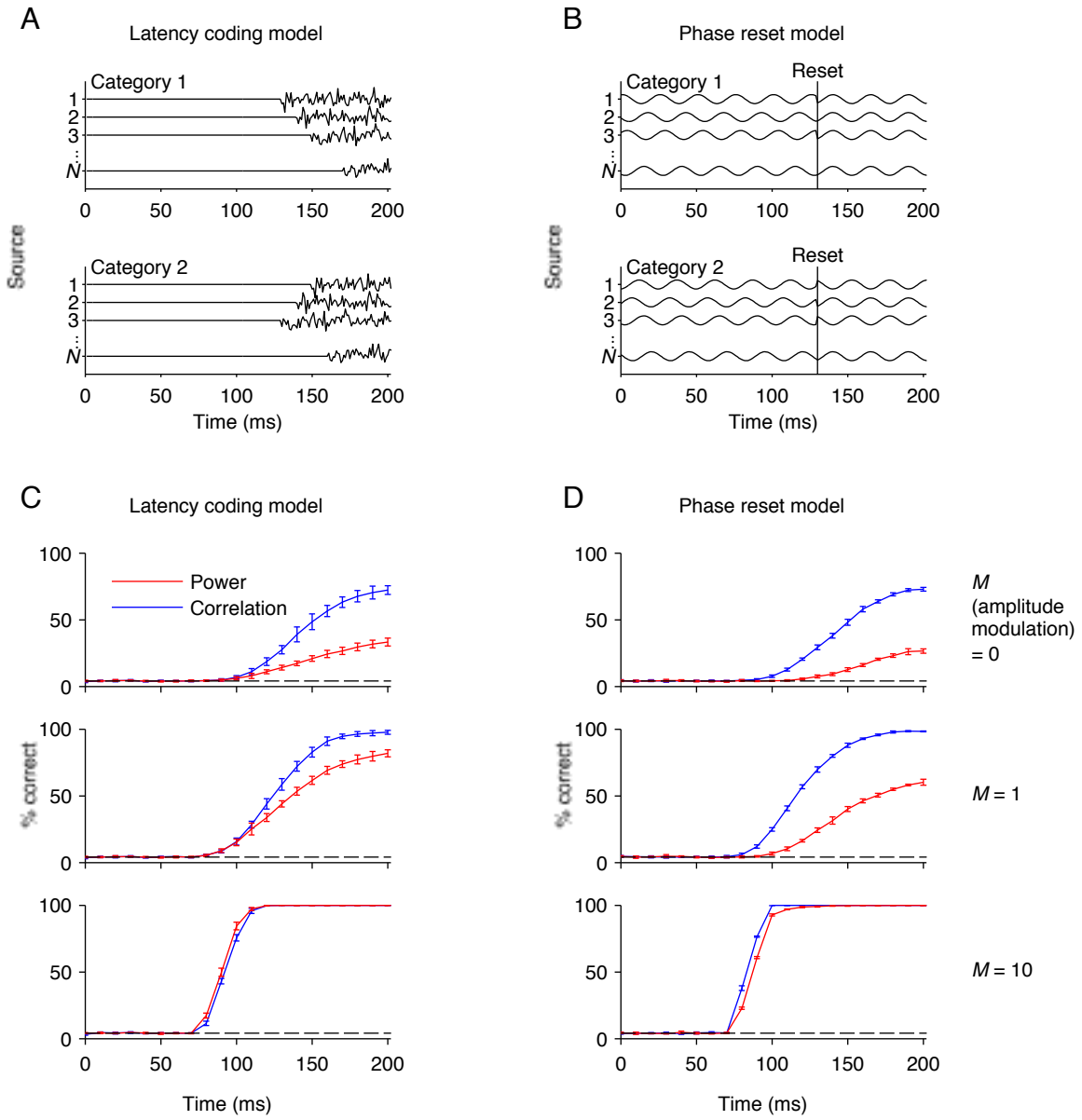

Figure 5.9. Simulation analysis. ***A,*** Neuronal responses in the latency-coding model (before adding noise). In each trial, a source produces its own waveform with a latency specific to the given stimulus. ***B,*** Neuronal responses in the phase-reset model (before adding noise). Sources behave as oscillators at a fixed frequency. In each trial, the phase of each source is reset to a value specific to the given stimulus at a time after the stimulus onset. ***C,*** Time course of decoding performance for the latency-coding model. The time courses of decoding performance with no amplitude modulation ($M = 0$, top), with moderate amplitude modulation ($M = 1$, middle), and with strong amplitude modulation ($M = 10$, bottom) are shown. ***D,*** Time course of decoding performance for the phase-reset model. The time courses of decoding performance with the three levels of amplitude modulation are shown.

5.4 Discussion

In the present chapter, we have shown that temporal correlations of ECoG signals between electrodes provide additional information on visual object categories compared to information represented by spectral power or phase in individual electrodes (Figs. 5.5 and 5.6). The trial-shuffling procedure degraded decoding performance obtained with correlations for raw and gamma band ECoG signals (Fig. 5.7), indicating that relative time series between electrodes contain information about categories. Time course analysis using a sliding time window revealed that decoding performance obtained with correlations began to rise earlier compared with power (Fig. 5.8). In the simulation analysis, this difference between power and correlation features was reproduced in a model where neuronal electric sources encode information using their phase differences (Fig. 5.9).

These results suggest that not only response amplitudes but also temporal correlations over multiple brain areas carry information on visual object categories, and that informative correlations can be derived from category-specific, relative time series of neural activity. The simulation results suggest that temporal correlations may reflect phase differences of neuronal electric sources in the temporal cortex at least in the early response period.

Although we focused on correlations of ECoG time series between electrodes without a time lag, the use of different features that take into account interactions between two different points in space and time may improve decoding performance. In a previous EEG study, it was reported that estimated off-diagonal coefficients in a multivariate autoregressive model are more informative than diagonal ones in classifying sleep stages (Penny and Roberts, 2002). It remains to be seen whether such sophisticated features are efficient for the decoding of neural object representations.

Several previous studies have proposed the possibility that visual information is encoded in temporal relations of spikes or field potentials between multiple brain locations, and more specifically, in the order of spike latencies (Gautrais and Thorpe, 1998; Thorpe et al., 2001; Van Rullen et al., 1998), spike sequences with

millisecond precision (Abeles, 1991; Abeles et al., 1993; Lestienne and Strehler, 1987; Oram et al., 1999), and gamma band synchronization (Eckhorn et al., 1988; Engel et al., 1991; Gray et al., 1989). Some experimental results have shown that such patterns encode low-level visual information, such as light intensity in the retina (Gollisch and Meister, 2008), orientation in the primary visual cortex (Celebrini et al., 1993; Gawne et al., 1996; Shriki et al., 2012), and co-occurrence of edges (Eckhorn et al., 1988; Engel et al., 1991; Gray et al., 1989).

Although our experimental data are not at the level of single neurons, ECoG signals have been assumed to provide an aggregate signature of transmembrane currents within a cortical area of several millimeters (Buzsáki et al., 2012; Reimann et al., 2013; Varela et al., 2001). Thus, informative correlations across electrodes in the IT cortex shown in our present study may derive from such temporal patterns encoding visual object categories.

Our shuffling procedure degraded decoding performance obtained with correlation features for the raw ECoG signals and the signals from the two highest frequency bands (Fig. 5.7). These results suggest that informative correlations partly arise from relative time series that are specific to each category and not time-locked to the stimulus onset across trials.

On the other hand, the shuffling procedure did not degrade decoding performance in the three lower frequency bands. This may imply that informative correlations in those bands arise in a stimulus-locked manner. However, it is also possible that even if relative time series did contribute to information coding in the lower frequency bands, the fluctuations of response delays across trials were too small relative to the low frequency cycles to have a substantial effect from shuffling.

We found that decoding performance with temporal correlations increased earlier than that with power values, and that the difference in the timing of the performance increase was reproduced by our phase-reset model, where object categories were coded by phase differences rather than latency differences across neuronal sources. While latency differences can cause robust patterns in both power and

correlation features, phase resetting may produce distinct patterns in correlation features but not in power features. Even in the phase reset model, spectral power could have been modulated via the synchronization or desynchronization of neuronal sources near each electrode. However, the detection of such small changes in power in the presence of noise may require a broader temporal summation, resulting in a delayed increase in decoding performance. The results suggest that at least in the early response period, informative correlations originate from subtle temporal patterns such as phase differences.

6. General Discussion

In this thesis, to improve decoding technology with ECoG, the signal-stability of the newly-developed electrode array was tested, the utility of ECoG was confirmed in monkey and human experiments, and efficient input signal features for decoding analysis was explored through four studies.

In chapter 2, by analyzing ECoG responses recorded via the electrode-mesh from the visual cortex in rats, it was shown that statistically significant decoding performance was attainable across six hours of data, which support the view that the electrode-mesh is one of the promising candidates for the next-generation input devices for BMI.

In chapter 3, by analyzing simultaneously recorded ECoG, LFP and MUA responses from the IT cortex in monkeys, it was demonstrated that ECoG's classification performance for visual object categories was high and comparable with LFP and MUA. Although a few previous studies have conducted simultaneous recording and compared them (Toda et al., 2011; Watanabe et al., 2012), their comparison was neither fair in terms of the number of electrodes, nor the spatial coverage. Miyakawa and his coworkers succeeded in recording ECoG, LFP, and MUA signals with the same number of electrodes in the same IT region, and we reported the first direct comparison of decoding performance between these recording modalities.

In chapter 4, we demonstrated that ECoG could be used to investigate how face-selective areas and letterstring-selective regions are distributed in the human vOT, which is considered a challenging task with fMRI. Brain regions that are selectively activated by face images and written words have been discovered in the vOT with fMRI. However, how those regions are distributed within the vOT was not clear because fMRI responses in the vicinity of the bottom of a brain are very noisy due to considerable inhomogeneity of the blood oxygen level–dependent signals (Ojemann et al. 1997). To address this issue, we used ECoG. The standard electrode-wise analysis conducted by Matsuo et al. (2013) and the decoding analysis conducted here have revealed that there exist separate,

multiple face-selective and written word-selective regions on the human vOT.

In chapter 5, it was shown that a decoder's predictive performance could be improved by using temporal correlations between ECoG electrodes, which reflects neural coding with phase differences across brain regions. In previous studies, spectral powers and signal amplitudes were often highlighted and used as input signal features for decoding analysis. However, theoretical studies have proposed a possibility that synchronizations or latency/phase differences between neuronal signals from different points encode information. If such coding were adopted in a recorded region, the use of features that directly reflect latency/phase differences would improve a decoder's predictive performance.

Here, we used temporal correlations between electrodes as features, and found that the decoding performance obtained using temporal correlations was higher than that with spectral powers and phases in individual electrodes. Also, analysis with a sliding time window revealed that we could extract information earlier with temporal correlations, compared to spectral powers.

From the analyses conducted above, it can be concluded that the combination of ECoG recordings and neural decoding techniques is a powerful approach for extracting neural information, and we can considerably improve a decoder's predictive performance by using signal features that take into account fine temporal patterns in ECoG signals.

Value of ECoG

ECoG allows recording of electric signals simultaneously from multiple sites and at high temporal resolution. This advantage is likely to have contributed to the high decoding performances shown in this thesis, and was especially critical in the shuffling analysis in chapter 5. In the shuffling analysis, we examined whether informative correlations derived from category-specific relative time series in each trial. This kind of analysis requires simultaneous recording of neural signals from multiple brain locations with high temporal resolution. In some studies using single-unit recordings, the signal from each electrode was recorded

separately because of the difficulty of simultaneous recording from multiple sites (Georgopoulos and Massey, 1988; Gochin et al., 1994; Hung et al., 2005; Kiani et al., 2007; McAdams and Maunsell, 1999; Rolls et al., 1997). However, with this method, stimulus-unlocked patterns over multiple brain locations may be overlooked (Averbeck et al., 2006). Recording with a microelectrode array could be a promising method for such an analysis, but the coverage of current state-of-the-art *in vivo* arrays is limited to an area of several millimeters and is not sufficient to cover the whole temporal cortex (Fejtl et al., 2006).

Although other neuroimaging techniques such as fMRI, scalp EEG, and MEG provide simultaneous recordings from wide brain regions, the temporal resolution of fMRI is limited to the timescale of seconds, and the spatial resolution of scalp EEG and MEG is insufficient to reveal synchronization within two or less centimeters (Varela et al., 2001). ECoG provides better spatiotemporal resolution than EEG, MEG, and fMRI, thus making ECoG a useful tool to examine theoretical hypotheses about fine temporal coding. The combination of ECoG recordings and neural decoding techniques has become a powerful approach for extracting information (Chao et al., 2010; Hammer et al., 2013; Liu et al., 2009; Pasley et al., 2012; Tsuchiya et al., 2008; Yanagisawa et al., 2009).

In addition, new, high-density ECoG electrode arrays are being developed (Hollenberg et al., 2006; Matsuo et al., 2011; Rubehn et al., 2009; Toda et al., 2011; Watanabe et al., 2012; Yeager et al., 2008). The use of these electrode arrays in monkeys and other animals may reveal even more detailed neural representations in the future.

Another advantage of ECoG is its signal stability. Previous studies have shown that ECoG can provide signals stable across even several months (Chao et al., 2010; Shimoda et al., 2012), while MUA signals from inserted electrodes deteriorate in at least several hours (Chestek et al., 2007).

In addition, since our mesh-film probe in chapter 2 was very thin and flexible, it was less invasive than the conventional clinical ECoG sheet. Although further

clinical tests are needed to establish long-term stability for months and years, these advantages are essential for implementing chronic neuroprosthetic devices for real-life applications.

Decoding analysis using temporal correlations for the data in chapter 4

Finally, we show the results of applying the decoding analysis in chapter 5 to the data in chapter 4. In chapter 4, brain regions selective to faces and letterstrings were identified by electrode-wise analysis. However, in chapter 5, not only spectral powers in individual electrodes, but also temporal correlations between electrodes could be informative features for decoding. To examine whether informative correlations arose within the face-/letterstring-selective regions, or if combinations of face-/letterstring-selective electrodes and electrodes in other regions were more informative, decoding performances were calculated for when the temporal correlations of ECoG time series for pairs of the face-/letterstring-selective electrodes (within-zone electrode pairs) were used as input features for decoding. Ten input pairs were randomly chosen from all available pairs to match the numbers of input features for comparison.

For face classification, the decoding performances obtained with features for within-zone electrode pairs were higher than those using correlations for electrode pairs between the face-selective electrodes and electrodes selective to other images (across-zone electrode pairs) (Fig 6.1A; $p < 0.05$). In addition, to examine whether the correlations for the across-zone electrode pairs had additional information, we compared the decoding performances obtained from the within-zone electrode pairs to those from both the within- and across-zone pairs. However, a performance improvement from adding the across-zone pairs was not observed (Fig 6.1A). For letterstring classification, the temporal correlations for the electrode pairs between the letterstring-selective electrodes and electrodes selective to other images showed lower performances than for those with correlations for the within-zone electrode pairs (Fig 6.1B; $p < 0.05$), and did not carry additional information on the category.

These results suggest that face- and letterstring-selective electrodes defined by the electrode-wise analysis contained more information for faces and letterstrings, respectively, for decoding analysis using temporal correlations.

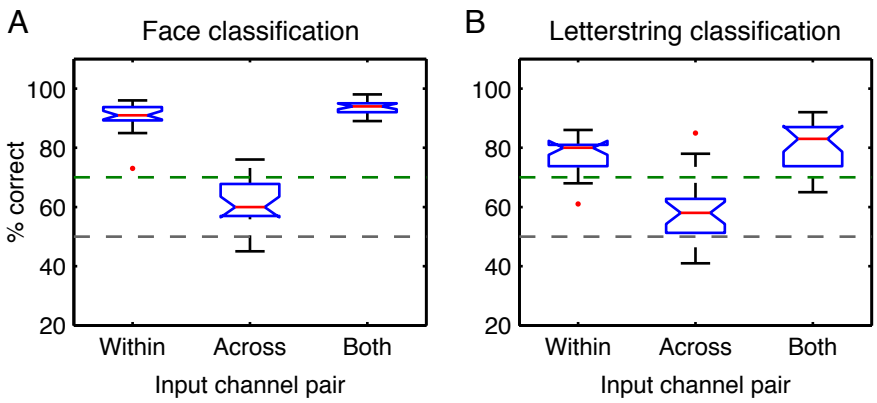

Figure 6.1. Binary classification performance obtained with temporal correlations between electrodes. ***A,*** Box plot of accuracies for discriminating faces from others. Classification accuracies were calculated using temporal correlations for pairs of face-selective electrodes (within), pairs between face-selective electrodes and others (across), and for both within and across electrodes (both). ***B,*** Box plot of accuracies in discriminating letterstrings from other images. Classification accuracies were calculated using temporal correlations for the pairs of the letterstring-selective electrodes (within), the pairs between the letterstring-selective electrodes and the others (across), and both of them (both). For both A and B, the boxplots signify the upper ($q3$) and lower ($q1$) quartiles, with the median depicted as the central red bar. The whiskers indicate the range and data points exceeding $q3 + 1.5 \times (q3 - q1)$ or below $q1 + 1.5 \times (q3 - q1)$ were treated as outliers and plotted individually as red dots; broken lines indicate chance level (gray) and the threshold of significance level (green, $p = 0.01$).

Acknowledgements

There are many people I would like to thank for their help and support with my research. First, I would like to offer my special thanks to my direct supervisor Yukiyasu Kamitani for his continuous support and greatly helpful advice on my work. For accepting me as a student at ATR, which is a favorable environment to focus on research work, I would like to thank Mitsuo Kawato. For helpful support and guidance as my supervisors, I would like to thank Kazushi Ikeda and Tomohiro Shibata. For being a member of my thesis committee, I would like to thank Kenji Sugimoto.

For helpful discussions and comments, I would like to thank members of student seminar at ATR and MIND meeting at DNI including Tomoyasu Horikawa, Makoto Fukushima, Makoto Takemiya, Yukina Hosokawa, Paul Sukhanov, Yusuke Fujiwara, Okito Yamashita, Kazuhisa Shibata, Ryohei Fukuma, Takufumi Yanagisawa, Keiji Harada, Mitsuaki Tsukamoto, and Carolin Lutz. Also, for scheduling those meetings, I would like to thank Natsuko Nakano.

For collecting and providing experimental data, I appreciate Isao Hasegawa, Takeshi Matsuo, Naohisa Miyakawa, Haruo Toda, Keisuke Kawasaki, Hisashi Tanigawa, Ken Adachi, Kiyoshi Nakahara, Atsuhiko Iijima and Hirohito Sawahata. For preparing experimental environments in The University of Tokyo Hospital and Nishi-Niigata Chuo National Hospital, I appreciate Kensuke Kawai, Hiroshi Masuda, Hiroatsu Murakami, Naoto Kunii, Shigeki Kameyama and Nobuhito Saito. For developing electrodes used in those experiments, I appreciate Takafumi Suzuki. Without their collaboration, all my work could not have been achieved.

I also would like to previous members who left ATR including Yoichi Miyawaki, Atsushi Matsui, Satoshi Murata, Masako Tamaki, Kentaro Yamada, Yasuhito Sawahata, Kaoru Amano, Yoriko Yamamura, Yohei Yamada, Megumi Fukuda, Hiromi Yamamura, Taylor Beck, Takatomi Kubo, Satoshi Hirose, Atsunori Kanemura, Yoshiyuki Onuki, Nancy Wang, Jason Bens, and previous members who left NAIST including Kohei Hayashi, Masayoshi Nakamura, Satoshi Nishida, Naoya Osugi, Atsushi Miyamoto, Satomi Adachi, Aya Tanimoto, Naoko Koide, Yuji

Yoshida, Satoshi Kozawa, Shigeyuki Ikeda, Tatsuya Yamada, Yuichi Sakumura, Satoshi Morimoto, Mauricio Burdelis, Takamitsu Araki, Ryota Aisu, Ryunosuke Hamada, Taiyo Kato.

I also would like to thank Takashi Sato, Hironobu Osaki, and Masashi Hasegawa for giving me a great chance to visit their laboratory in Tübingen. Although our collaboration in Tubingen is not directly related with this thesis, I am sure that what I have experienced there has a strong impact on my way of research.

In addition, I would like to thank all others who have supported me and given me the ability to make it this far.

References

[1] Abeles M (1991) Corticonics: Neural Circuits of the Cerebral Cortex. New York: Cambridge UP.

[2] Abeles M, Bergman H, Margalit E, Vaadia E (1993) Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. J Neurophysiol 70:1629-1638.

[3] Allison T, Puce A, Spencer D, McCarthy G (1999) Electrophysiological studies of human face perception. I: potentials generated in occipitotemporal cortex by face and non-face stimuli. Cereb Cortex 9:415-430.

[4] Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci 7:358-366.

[5] Bialek W, Rieke F, de Ruyter van Steveninck RR, Warland D (1991) Reading a neural code. Science 252:1854-1857.

[6] Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13: 407-420.

[7] Celebrini S, Thorpe S, Trotter Y, Imbert M (1993) Dynamics of orientation coding in area V1 of the awake primate. Vis Neurosci. 10:811-825.

[8] Chang CC, Lin CJ. (2011) LIBSVM: A library for support vector machines. Acm T Intel Syst Tec 2:27.

[9] Chao ZC, Nagasaka Y, Fujii N (2010) Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys. Front Neuroeng 30:3.

[10] Chestek CA, Batista AP, Santhanam G, Yu BM, Afshar A, Cunningham JP, Gilja V, Ryu SI, Churchland MM, Shenoy KV (2007) Single-neuron stability during repeated reaching in macaque premotor cortex. J Neurosci 27:10742-10750.

[11] Cohen L, Martinaud O, Lemer C, Lehericy S, Samson Y, Obadia M, Slachevsky A, Dehaene S (2003) Visual word recognition in the left and right hemispheres: anatomical and functional correlates of peripheral alexias. Cereb Cortex 13:1313-1333.

[12] Dayan P, Abbott, LF (2001) Theoretical neuroscience. Cambridge MA: MIT Press.

[13] Dehaene S, Cohen L (2011) The unique role of the visual word form area in reading. Trends Cogn Sci 15:254-262.

[14] Donoghue JP (2002) Connecting cortex to machines: recent advances in brain interfaces. Nat Neurosci 5: 1085-1088.

[15] Eckhorn R, Bauer R, Jordan W, Brosch M, Kruse W, Munk M, Reitboeck HJ (1988) Coherent oscillations: a mechanism of feature linking in the visual cortex? Multiple electrode and correlation analyses in the cat. Biol Cybern 60:121-130.

[16] Engel AK, König P, Kreiter AK, Singer W (1991) Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. Science 252:1177-1179.

[17] Fejtl M, Stett A, Nisch W, Boven KH, Möller A (2006) On micro-electrode array revival: its development, sophistication of recording, and stimulation. In: Advances in Network Electrophysiology: Using Multi-electrode Arrays (Taketani M, Baudry M, ed), pp24-37. New York: Springer.

[18] Fisch L, Privman E, Ramot M, Harel M, Nir Y, Kipervasser S, Andelman F, Neufeld MY, Kramer U, Fried I, Malach R (2009): Neural "ignition": enhanced activation linked to perceptual awareness in human ventral stream visual cortex. Neuron 64:562-574.

[19] Gautrais J, Thorpe S (1998) Rate coding versus temporal order coding: a theoretical approach. Biosystems 48:57-65.

[20] Gawne TJ, Kjaer TW, Richmond BJ (1996) Latency: another potential code for feature binding in striate cortex. J Neurophysiol 76:1356-1360.

[21] Georgopoulos AP, Schwartz AB, Kettner RE (1986) Neuronal population coding of movement direction. Science 233:1416-1419.

[22] Georgopoulos AP, Massey JT (1988) Cognitive spatial-motor processes. 2. Information transmitted by the direction of two-dimensional arm movements and by neuronal populations in primate motor cortex and area 5. Exp Brain Res 69:315-326.

[23] Girman SV, Sauve Y, Lund RD, (1999) Receptive field properties of single neurons in rat primary visual cortex. J Neurophysiol 82: 301-311.

[24] Gochin PM, Colombo M, Dorfman GA, Gerstein GL, Gross CG (1994) Neural ensemble coding in inferior temporal cortex. Neurophysiol 71:2325-2337.

[25] Gray CM, König P, Engel AK, Singer W (1989) Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. Nature 338:334-337.

[26] Gollisch T, Meister M (2008) Rapid neural coding in the retina with relative spike latencies. Science 319:1108-1111.

[27] Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. Nature 458:632-635.

[28] Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425-2430.

[29] Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD, Donoghue JP (2006) Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 442:164-171.

[30] Hollenberg BA, Richards CD, Richards R, Bahr DF, Rector DM (2006) A MEMS fabricated flexible electrode array for recording surface field potentials. J Neurosci Methods 153:147-153.

[31] Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y (2013) Neural decoding of visual imagery during sleep. Science 340:639-642.

[32] Hung CP, Kreiman G, Poggio T, DiCarlo JJ (2005) Fast readout of object identity from macaque inferior temporal cortex. Science 310:863-866.

[33] Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. Nat Neurosci 8:679-685.

[34] Kamitani Y, Tong F (2006) Decoding seen and attended motion directions from activity in the human visual cortex. Curr Biol 16:1096-1102.

[35] Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17:4302-4311.

[36] Kanwisher N, Yovel G (2006) The fusiform face area: a cortical region specialized for the perception of faces. Philos Trans R Soc Lond B Biol Sci 361:2109-2128.

[37] Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J Neurophysiol 97:4296-4309.

[38] Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12:535-540.

[39] Lebedev MA, Nicolelis MA. (2006) Brain-machine interfaces: past, present and future. rends Neurosci 29:536-546.

[40] Lesaffre E, Rizopoulos D, Tsonaka R (2007) The logistic transform for bounded outcome scores. Biostatistics 8:72-85.

[41] Lestienne R, Strehler BL (1987) Time structure and stimulus dependence of precisely replicating patterns present in monkey cortical neuronal spike trains. Brain Res 437:214-238.

[42] Leuthardt EC, Schalk G, Wolpaw JR, Ojemann JG, Moran DW (2004) A brain-computer interface using electrocorticographic signals in humans. J Neural Eng 1:63-71.

[43] Leuthardt EC, Schalk G, Moran D, Ojemann JG (2006) The emerging world of motor neuroprosthetics: a neurosurgical perspective. Neurosurgery 59:1-14.

[44] Liu H, Agam Y, Madsen JR, Kreiman G (2009) Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. Neuron 62: 281-290.

[45] Majima K, Matsuo T, Kawasaki K, Kawai K, Saito N, Hasegawa I, Kamitani Y (in press) Decoding visual object categories from temporal correlations of ECoG signals. Neuroimage.

[46] Matsuo T, Kawasaki K, Osada T, Sawahata H, Suzuki T, Shibata M, Miyakawa N, Nakahara K, Iijima A, Sato N, Kawai K, Saito N, Hasegawa I (2011) Intrasulcal electrocorticography in macaque monkeys with minimally invasive neurosurgical protocols. Front Syst Neurosci 5:34.

[47] Matsuo T, Kawasaki K, Kawai K, Majima K, Masuda H, Murakami H, Kunii N, Kamitani Y, Kameyama S, Saito N, Hasegawa I (in press) Alternating zones selective to faces and written words in the human ventral occipitotemporal cortex. Cereb Cortex.

[48] Miyakawa N, Majima K, Sawahata H, Kawasaki K, Kotake N, Matsuo T, Suzuki T, Kamitani Y, Hasegawa I (2011) Category representation in macaque anterior inferotemporal cortex revealed with simultaneous electrocorticogram and multi-channel unit recording. Soc Neurosci Abstr: 324.08.

[49] Miyakawa N, Majima K, Sawahata H, Kawasaki K, Matsuo T, Kotake N, Suzuki T, Kamitani Y, Hasegawa I (in preparation) Distinct sizes of neuronal clustering code hierarchically different-level categories in macaque anterior inferotemporal cortex.

[50] Miyawaki Y, Uchida H, Yamashita O, Sato MA, Morito Y, Tanabe HC, Sadato N, Kamitani Y (2008) Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. Neuron 60:915-929.

[51] Naruse Y, Matani A, Miyawaki Y, Okada M (2010) Influence of coherence between multiple cortical columns on alpha rhythm: a computational modeling study. Hum Brain Mapp 31:703-715.

[52] Naselaris T, Prenger RJ, Kay KN, Oliver M, Gallant JL (2009) Bayesian reconstruction of natural images from human brain activity. Neuron 63:902-915.

[53] Nishimoto S, Vu AT, Naselaris T, Benjamini Y, Yu B, Gallant JL (2011) Reconstructing visual experiences from brain activity evoked by natural movies.CurrBiol 21: 1641-1646.

[54] Nobre A, Allison T, McCarthy G (1994) Word recognition in the human inferior temporal lobe. Nature 372:260-263.

[55] Nunez PL (1981) Electric Fields of the Brain. The Neurophysics of EEG. New York: Oxford UP.

[56] Ojemann J, Akbudak E, Snyder A, McKinstry R, Raichle M, Conturo T (1997) Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. Neuroimage 6:156-167.

[57] Oram MW, Wiener MC, Lestienne R, Richmond BJ (1999) Stochastic nature of precisely timed spike patterns in visual system neuronal responses. J Neurophysiol 81:3021-3033.

[58] Palva S, Palva JM (2007) New vistas for alpha-frequency band oscillation. Trends Neurosci 30:150-158.

[59] Parvizi J, Jacques C, Foster B, Witthoft N, Withoft N, Rangarajan V, Weiner K, Grill-Spector K (2012) Electrical stimulation of human fusiform face-selective regions distorts face perception. J Neurosci 32:14915-14920.

[60] Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, Knight RT, Chang EF (2012) Reconstructing speech from human auditory cortex. PLoS Biol 10: e1001251.

[61] Paxinos G, Watson C (2007) The rat brain in stereotaxic coordinates, 6th ed. Academic Press, London.

[62] Penny WD, Roberts SJ (2002) Bayesian multivariate autoregressive models with structured priors. IEE Proc.-Vis Image Signal Process 149:33-41.

[63] Pistohl T, Ball T, Schulze-Bonhage A, Aertsen A, Mehring C (2008) Prediction of arm movement trajectories from ECoG-recordings in humans. J Neurosci Methods 167:105-114.

[64] Price C (2012) A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage 62:816-847.

[65] Reddy L, Tsuchiya N, Serre T (2010) Reading the mind's eye: decoding category information during mental imagery. Neuroimage 50:818-825.

[66] Reimann MW, Anastassiou CA, Perin R, Hill SL, Markram H, Koch C (2013) A biophysically detailed model of neocortical local field potentials predicts the critical role of active membrane currents. Neuron 79:375-390.

[67] Rolls ET, Treves A, Tovee MJ (1997) The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. Exp Brain Res 114:149-162.

[68] Rubehn B, Bosman C, Oostenveld R, Fries P, Stieglitz T (2009) A MEMS-based flexible multichannel ECoG-electrode array. J Neural Eng 6:036003.

[69] Schalk G, Miller KJ, Anderson NR, Wilson JA, Smyth MD, Ojemann JG, Moran DW, Wolpaw JR, Leuthardt EC (2008) Two-dimensional movement control using electrocorticographic signals in humans. J Neural Eng 5:75-84.

[70] Serruya MD, Hatsopoulos NG, Paninski L, Fellows MR, Donoghue JP (2002) Instant neural control of a movement signal. Nature 416:141-142.

[71] Shimoda K, Nagasaka Y, Chao ZC, Fujii N (2012) Decoding continuous three-dimensional hand trajectories from epidural electrocorticographic signals in Japanese macaques. J Neural Eng 9: 036015.

[72] Shriki O, Kohn A, Shamir M (2012) Fast coding of orientation in primary visual cortex. PLoS Comput Biol 8:e1002536.

[73] Stanley GB (2013) Reading and writing the neural code. Nat Neurosci 16:259-263.

[74] Takeuchi S, Ziegler D, Yoshida Y, Mabuchi K, Suzuki T (2005) Parylene flexible neural probes integrated with microfluidic channels. Lab Chip 7:519-523.

[75] Talairach J, Tournoux P (1993) Referentially Oriented Cerebral MRI Anatomy: An Atlas of Stereotaxic Anatomical Correlations for Gray and White Matter. New York: Thieme Medical Publishers.

[76] Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 3:151-162.

[77] Tallon-Baudry C, Bertrand O, Pernier J (1999) A ring-shaped distribution of dipoles as a source model of induced gamma-band activity. Clin Neurophysiol 110:660-665.

[78] Thorpe S, Delorme A, Van Rullen R (2001) Spike-based strategies for rapid processing. Neural Netw 14:715-725.

[79] Toda H, Suzuki T, Sawahata H, Majima K, Kamitani Y, Hasegawa I (2011) Simultaneous recording of ECoG and intracortical neuronal activity using a flexible multichannel electrode-mesh in visual cortex. Neuroimage 54:203-212.

[80] Truccolo W, Friehs GM, Donoghue JP, Hochberg LR (2008) Primary motor cortex tuning to intended movement kinematics in humans with tetraplegia. J Neurosci 28:1163-1178.

[81] Tsuchiya N, Kawasaki H, Oya H, Howard MA 3rd, Adolphs R (2008) Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. PLoS One 3:e3892.

[82] Van Rullen R, Gautrais J, Delorme A, Thorpe S (1998) Face processing using one spike per neurone. Biosystems 48:229-239.

[83] Vapnik V (1998) Statistical Learning Theory. New York: John Wiley & Sons.

[84] Varela F, Lachaux JP, Rodriguez E, Martinerie J (2001) The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci 2:229-239.

[85] Vindiola M, Wolmetz M (2011) Mental encoding and neural decoding of abstract cognitive categories: a commentary and simulation. Neuroimage 54:2822-2827.

[86] Watanabe H, Sato MA, Suzuki T, Nambu A, Nishimura Y, Kawato M, Isa T (2012) Reconstruction of movement-related intracortical activity from micro-electrocorticogram array signals in monkey primary motor cortex. J Neural Eng 9:036006.

[87] Wessberg J, Stambaugh CR, Kralik JD, Beck PD, Laubach M, Chapin JK, Kim J, Biggs SJ, Srinivasan MA, Nicolelis MA (2000) Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. Nature 408:361-365.

[88] Winawer J, Horiguchi H, Sayres R, Amano K, Wandell B (2010) Mapping hV4 and ventral occipital cortex: the venous eclipse. J Vis 10:1.

[89] Yamamura H, Sawahata Y, Yamamoto M, Kamitani Y (2009) Neural art appraisal of painter: Dali or Picasso? Neuroreport 20:1630-1633.

[90] Yamashita O, Sato MA, Yoshioka T, Tong F, Kamitani Y (2008) Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. Neuroimage 42:1414-1429.

[91] Yanagisawa T, Hirata M, Saitoh Y, Kato A, Shibuya D, Kamitani Y, Yoshimine T (2009) Neural decoding using gyral and intrasulcal electrocorticograms. Neuroimage 45:1099-1106.

[92] Yeager JD, Phillips DJ, Rector DM, Bahr DF (2008) Characterization of flexible ECoG electrode arrays for chronic recording in awake rats. J Neurosci Methods 173:279-285.

[93] Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY (2008) Transient induced gamma-band response in EEG as a manifestation of miniature saccades. Neuron 58:429-441.