33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Decoding Semantic Information from Human Electrocorticographic (ECoG) Signals

Wei Wang, Alan D Degenhart, Gustavo P Sudre, Dean A Pomerleau, Elizabeth C. Tyler-Kabara

Abstract—**This study examined the feasibility of decoding semantic information from human cortical activity. Four human subjects undergoing presurgical brain mapping and seizure foci localization participated in this study. Electrocorticographic (ECoG) signals were recorded while the subjects performed simple language tasks involving semantic information processing, such as a picture naming task where subjects named pictures of objects belonging to different semantic categories. Robust high-gamma band (60-120Hz) activation was observed at the left inferior frontal gyrus (LIFG) and the posterior portion of the superior temporal gyrus (pSTG) with a temporal sequence corresponding to speech production and perception. Furthermore, Gaussian Naïve Bayes and Support Vector Machine classifiers, two commonly used machine learning algorithms for pattern recognition, were able to predict the semantic category of an object using cortical activity captured by ECoG electrodes covering the frontal, temporal and parietal cortices. These findings have implications for both basic neuroscience research and development of semantic-based brain-computer interface systems (BCI) that can help individuals with severe motor or communication disorders to express their intention and thoughts.**

I. Introduction

Brain-computer interface (BCI) technology aims to help individuals with disabilities by establishing a direct link between the brain and external devices, enabling faster and more intuitive communication and control [1]. There has been a large of number of BCI studies focusing on extracting movement signals from motor cortical activity in order to control a computer cursor or a prosthetic limb [2, 3]. For a BCI system to serve as an assistive device to enable individuals with communication disorders to express their intention, it is potentially desirable to directly extract the meaning, i.e. the semantic information an individual is thinking about, from cortical activity.

Semantic information, also referred to as conceptual knowledge or semantic memory, is essentially the concept of a specific object (e.g. an apple or car) or action (e.g. grasp or kick) [4]. Cognitive neuroscience studies using functional magnetic resonance imaging (fMRI) have suggested that there exists a distributed representation of semantic information in the cerebral cortex [5]. One recent fMRI study showed that the semantic category (e.g. tools vs. buildings) of an object that an individual is thinking about can be predicted from blood-oxygen-level dependence (BOLD) signals, suggesting the feasibility of decoding semantic information from cortical activity [6]. While fMRI has been instrumental in the exploration of higher brain functions in humans, its application in BCI is potentially limited due to its relatively low temporal resolution and lack of portability. Recently, electrocorticography (ECoG) has been demonstrated as a promising cortical recording technique for BCI applications [7-9]. ECoG records brain activity with intracranial electrodes placed directly on the brain surface, offering high signal-to-noise ratio, high spatiotemporal resolution, and potentially broad coverage of multiple cortical networks. It has already received considerable attention as a practical modality for extracting motor control signals (e.g. hand and finger movement signals) from cortical activity for BCI applications [7, 9]. Combining previous fMRI findings and ECoG technology, the current study aims to examine the feasibility of an ECoG-based BCI system that operates by extracting semantic information from human cortical activity.

Manuscript received June 20, 2011. This work was supported by US Army Medical Research and Material Command Agreement W81XWH-07-1-0716, the Craig H. Neilsen Foundation, NIH Grants 5 UL1 RR024153 and KL2 RR024154, and NIH/NINDS grant 3R01NS050256-05S1.

Wei Wang, Alan Degenhart, Elizabeth Tyler-Kabara are with the University of Pittsburgh, Departments of Physical Medicine and Rehabilitation, Bioengineering, and Neurological Surgery, respectively, PA 15213, USA (e-mail: wangwei3@pitt.edu,elizabeth.tyler-kabara@chp.edu).

Gustavo Sudre is with the Carnegie Mellon University Program in Neural Computation, PA 15213, USA.

Dean Pomerleau is with Intel Labs Pittsburgh and Carnegie Mellon University, PA 15213, USA.

II. Methods

A. Human subjects and behavioral paradigms

This study was approved by the local Institutional Review Board and followed all guidelines for human subject research. Informed consent was obtained for all subjects before the initiation of any research procedures. Data presented in this paper were collected from four right-handed subjects undergoing clinical ECoG monitoring for intractable epilepsy (Table 1). Subjects performed a baseline task and three language tasks. During the baseline task, the subject was instructed to relax with eyes open for one minute. For all language tasks, pictures of various objects belonging to different semantic categories were presented to the subject on a computer screen in a pseudo-random order (one object per trial). For each trial, the picture stayed on for 2~4 seconds determined randomly by the computer. The inter-trial-interval between picture presentations was 2 seconds. Table 1 lists the language task performed by each subject during each experiment session. The "picture naming" task was performed by all four subjects, and the instruction was to simply name the object. Additionally, Subject A performed a "property identification" task, where she was instructed to think about various sensory and motor properties of the object throughout a trial, similar to tasks used in previous fMRI studies [6]. Subject B also performed a "closest word" task, where she was instructed to name an object that was closest to the one shown in the picture. Objects for Subject A were drawn from the food, tool, and dwelling categories, while objects for Subjects B, C, and D were drawn from the food, tool, and body part categories. The "property identification" task used fifteen objects with five repetitions each. The other two language tasks used eight objects with six repetitions each.

B. System setup and ECoG recording

Standard clinical ECoG grids (Ad-Tech Corp., Racine, WI, USA) with between 48 and 108 electrodes (3 mm diameter contact area, 10 mm center-to-center distance) were implanted subdurally over the left hemisphere for epilepsy seizure monitoring and presurgical brain mapping. In addition, for Subjects B, C, and D, a high-density research grid was implanted (Ad-Tech Corp.). This research grid had 16 disc electrodes (1.5 mm diameter contact area, 4 mm center-to-center distance) [10]. Figure 1 illustrates the placement of ECoG electrodes on a standard Montreal Neuroscience Institute (MNI) brain template [11]. All ECoG signals were band-pass filtered between 0.1 and 200 Hz and sampled at 1200 Hz using the g.USBamp neural recording system (Guger Technologies, Austria) in conjunction with the BCI2000 software package. The latter also controlled experiment paradigms and presented visual stimuli [12].

C. Neural data processing and decoding analysis

For both the baseline and language tasks, time-frequency distributions were calculated from the raw time domain signals using the maximum entropy method with a sliding window (window and step sizes of 300 and 50 ms, 1 Hz frequency bins) [13]. For each subject, the time-frequency distributions for language tasks were then log-transformed and converted to pseudo Z-scores using the following equation [14]:

$$r = (p - b_{mean}) / b_{std}$$

Here r is the pseudo Z-score, p is the log-transformed power of a specific 1-Hz wide frequency band, and b_{mean} and b_{std} are the mean and standard deviation of the log-transformed power for the same frequency band during the one-minute baseline task for the same subject. Pseudo Z-score data were then used in all of the subsequent data analysis.

Two types of analysis were performed. First, the high-gamma band activity (60-120 Hz), which was shown previously to be highly correlated with local neuronal activity during motor and cognitive tasks [15, 16], was examined as a function of time. Second, decoding analysis was performed to classify the semantic category of the presented object from ECoG signals recorded from all electrodes. ECoG data corresponding to objects of the same semantic category were grouped and averaged within the one second window after picture onset. Then, both Gaussian Naïve Bayes (GNB) and multi-class linear Support Vector Machine (SVM) classifiers [17] were used to separately decode semantic categories from single-trial cortical activity recorded from all ECoG electrodes with leave-one-out cross validation.

Subject	*Age*	*Gender*	*Electrodes*	*Session*	*Paradigm*	*Semantic categories*
A	12	Female	64	1	Property identification	Food/tools/dwellings
				2	Picture naming	Food/tools/dwellings
B	17	Female	64	1	Picture naming	Food/tools/body parts
				2	Closest word	Food/tools/body parts
C	23	Female	64	1	Picture naming	Food/tools/body parts
				2	Picture naming	Food/tools/body parts
D	12	Female	124	1	Picture naming	Food/tools/body parts

Table 1. Subject information and the language task performed by each subject during each session.

III. Results

A. Subject information and electrode placement

Table 1 lists the demographic information, total number of ECoG electrodes, and the language tasks performed by each subject. Figure 1 shows locations of ECoG electrodes, which were extracted from post-implant head x-ray images (lateral view) and mapped to the MNI brain template. For each subject, electrode locations were determined based on clinical requirements for his or her ECoG monitoring.

B. High-gamma band activity during the picture naming task

Figure 2 shows the spectrograms (i.e. time-frequency distributions) of ECoG signals recorded from two electrodes implanted in Subject A. The color represents the pseudo Z-score. These two electrodes were above the left inferior frontal gyrus (LIFG, traditionally considered Broca's area) and the posterior portion of the superior temporal gyrus (pSTG, traditionally considered Wernicke's area). These two areas are the essential cortical structures for language and speech functions. For both electrodes, the high-gamma band, which typically reflects local ensemble neuronal activity, showed significant increase in power during the picture naming task. To further illustrate the temporal dynamics of high-gamma band activity, the pseudo Z-score averaged across 60-120 Hz and across all objects was plotted as a function of time (Figure 3). The high-gamma band activity recorded from LIFG peaked at 0.75 seconds from picture onset, which is temporally correlated with speech production.

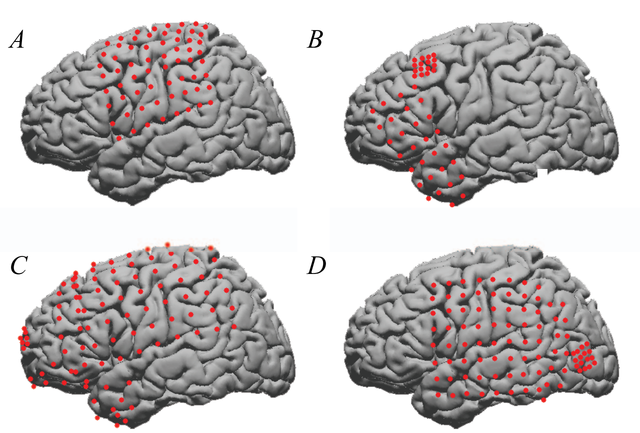

Figure 1. Electrode locations for Subjects A to D. Electrode locations have been mapped to a standard MNI brain template. Red dots represent implanted ECoG electrodes used in this study.

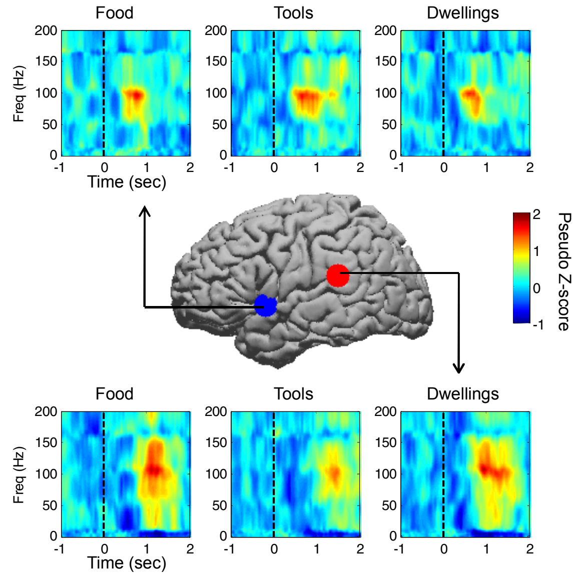

Figure 2. Time-frequency plots for ECoG signals recorded from electrodes over LIFG (blue dots and top panel) and pSTG (red dot and bottom panel) in Subject A. Time 0 indicates picture onset. Within the top and bottom panels, there are three plots corresponding to three semantic categories. Each plot shows data averaged across all objects of the same semantic category. Red colors indicate increases in spectral power relative to baseline, while blue colors indicate decreases relative to baseline.

The high-gamma band activity recorded from pSTG peaked at 1.2 seconds, when the subject perceived her own speech after naming an object. This result demonstrates the high spatial and temporal resolution of ECoG for capturing human cortical activity when subjects were actively processing semantic information in a language task.

C. *Decoding semantic categories from cortical activity*

Figure 4 shows accuracies obtained when predicting an object's semantic category from the three possible categories using both the GNB and SVM methods for each subject and each experiment session across all three language tasks. Both GNB and SVM were able to extract semantic categories from human cortical activity with accuracies as high as 74% (Chance level 33%). For all seven experiment sessions, SVM typically performed better than GNB. Furthermore, although statistical analysis was difficult to conduct with a limited number of experiment sessions, the decoding accuracy does not seem to differ significantly among the three different types of language tasks (picture naming, property identification, and closest word).

IV. Discussion

This study recruited four human subjects undergoing presurgical brain mapping and seizure foci localization. ECoG signals were recorded while subjects performed language tasks. High-gamma band activation was observed in classic language areas, LIFP and pSTG, and semantic information was extracted from neural signals record from ECoG electrodes spanning multiple cortical networks. This study has multiple implications for both cognitive

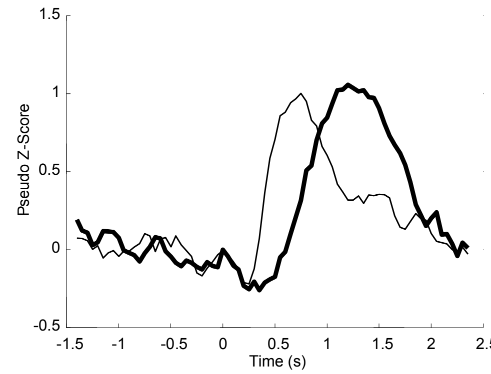

Figure 3. Temporal dynamics of high-gamma activity related to speech generation and perception during the picture-naming task (Subject A). The thin and thick lines represent the high-gamma band activities recorded from the electrodes over LIFG and pSTG (as shown in Figure 2), respectively. The high-gamma band activity was averaged over 60-120Hz and over all objects and all semantic categories.

neuroscience research and development of future brain-computer interface devices.

A. *Implications for cognitive neuroscience research*

In accordance with previous ECoG studies [14, 15], this study further demonstrates the uniqueness and potential of ECoG for studying human cortical processes related to speech and language, complementing non-invasive functional imaging techniques, such as fMRI and EEG. Since ECoG directly records electrical field potentials generated by neuronal activity, it has high temporal resolution, enabling researchers to examine cortical activity over a time scale (1~100 milliseconds) comparable to the neural processes of speech and language. Since ECoG electrodes are placed directly on the brain surface, it has high signal-to-noise ratio, which is specifically beneficial for capturing high-gamma band activity.

The sequential activation of LIFG and pSTG during the picture naming task as reflected by high-gamma band activity is similar to the findings of recent ECoG studies [14]. This supports the more traditional view that LIFG is engaged in speech production and pSTG is engaged in speech perception [18]. The most interesting finding from this study is that semantic information can be directly decoded from ECoG signals recorded during language tasks. In the previous fMRI study, the participants were consciously thinking about various sensory and motor properties of an object for several seconds while BOLD signals were measured [6]. The picture naming paradigm used in the current study has a fast pace comparable to natural speech behavior during daily living. Thus, the current study further extends the fMRI work and suggests that specific semantic information is embedded in the activity pattern of a distributed cortical network and that such semantic representation is innate to natural cortical processes for speech production and comprehension and is not caused by subjects' conscious effort in actively thinking about object properties.

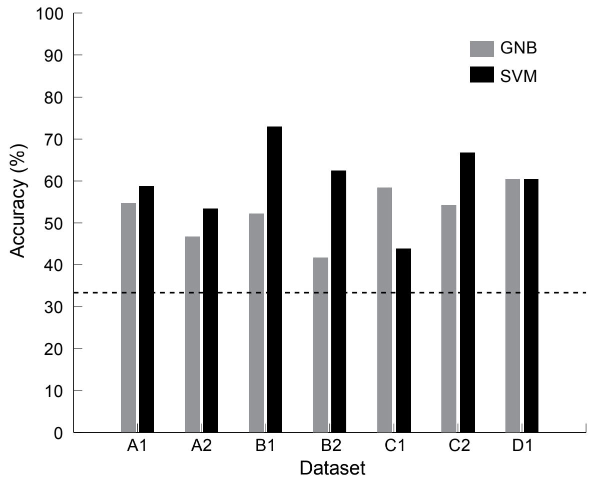

Figure 4. Classification accuracies for all datasets during the language tasks for both Gaussian Naïve Bayes (GNB) and multiclass linear Support Vector Machine (SVM) classifiers. Each pair of gray (GNB) and black (SVM) bars represents decoding accuracies for a specific dataset. Each dataset was labeled on the x-axis using Subject ID and session number as listed in Table 1. For example, "A1" indicates the dataset collected from Subject A's first session. Shown accuracies are obtained using ECoG time-frequency distribution data over the one-second post-picture-onset time window providing maximum classification accuracy. Chance accuracy (33.3%) is shown by the dashed line.

B. Implications for brain-computer interfaces

To date, a large of number of BCI studies have focused on extracting basic movement signals (e.g. position and velocity) in order to control a computer cursor or a robotic arm [2, 19, 20]. While these motor-based BCI systems are effective for controlling wheelchairs and prosthetic limbs, they may not be the most efficient when acting as a communication aid. Researchers have been exploring possibilities of developing BCI systems based on high-level cognitive functions for communication augmentation. One well-established example is the P300-speller, which utilizes the classic "odd ball" paradigm, where the item of interest to a user occurs less frequently than irrelevant items, thus evoking a distinct cortical response that identifies the target item [21]. As another example, researches have demonstrated the feasibility of extracting high-level cognitive signals, such as the goal and expected reward value from parietal and premotor areas to drive BCI systems [22]. Most recently, it has been proposed that the dorsolateral prefrontal cortical activity associated with working memory function can also be used to extract BCI control signals [23]. The current study focuses on cortical processes related to human speech and language functions, specifically semantic information processing. The study offers the initial evidence that semantic information can be extracted from human cortical activities. As the demonstration of decoding movement information from motor cortex has led to the explosive development of motor-based BCI systems [3, 7, 9, 19], we expect that our findings about semantic information decoding will inspire the development of semantic-based BCI as communication aids. Such systems may potentially serve as a natural, intuitive, and fast interface for an individual with severe disability to communicate with others.

C. Limitations and future directions

While results from the current study are encouraging, more thorough data analysis and collection of new datasets are warranted. For example, a detailed and quantitative examination of the neural decoders (GNB and SVM) will be needed to identify specific cortical areas that are the most informative for semantic information decoding. Comparison between results from such analysis and existing neuroimaging literature will further verify the basic findings from this study and also provide additional insights into cortical representation of semantic information. For example, it has been debated whether semantic information is simply represented in a distributed fashion or there is one or a few hubs, e.g. LIFG and the anterior temporal lobe (ATL), that specifically integrate activities of multiple cortical networks into a localized amodal abstract representation [4]. Future ECoG studies using semantic tasks might be able to provide new evidence to answer such questions.

The above basic neuroscience question also has implications for the development of a practical semantic-based BCI system using ECoG. If the representation is widely distributed across the cerebral cortex, a large grid of ECoG electrodes needs to be implanted in order to cover a significant portion of the cerebral cortex. While large ECoG grids are often used in clinical epilepsy monitoring, it will be impractical to implant such large grids for BCI applications. Alternatively, if there are central hubs for semantic information representation, such as LIFG or ATL, it will be more convenient to implant a small high-density electrode arrays to those hubs and extract semantic information.

Decoding methods used in this study also need to be further improved. The current approach does not specifically account for potentially correlated activities across multiple frequency bands recorded from the same ECoG electrode [24]. Therefore, it is expected that application of appropriate feature selection and dimensionality reduction techniques will further improve the decoding accuracy [6]. Additionally, decoding analysis used ECoG frequency band activity averaged over the one-second time window after picture onset. This approach does not take full advantage of the high temporal resolution provided by ECoG. Detailed temporal profiles of ECoG signals are potentially very informative as shown in Figure 3, and decoding analysis using full time-frequency distribution data is worth conducting. Lastly, as shown in Table 1, there are biases in age and gender given the small subject number. We plan to accrue more subjects to overcome such biases, but it is also worth noting that accrual rate for ECoG subjects will be slower than for healthy volunteers, as it depends on clinical flow of ECoG cases in affiliated medical centers. Finally, age, gender, and potential cortical changes due to epilepsy need to be taken into account for the generalization of results presented here to a larger population. In summary, this paper demonstrates successful decoding of semantic information from human ECoG signals, and this finding may pave the way for developing semantic-based BCI systems in the near future.

Acknowledgment

We would like to thank the participants who made it possible for us to perform this study. We would also like to thank all the clinical staff of epilepsy monitoring units. Dr. Tom Mitchell also provided insightful suggestions and comments.

References

[1] A.B. Schwartz, X.T. Cui, D.J. Weber, and D.W. Moran, "Brain-controlled interfaces: movement restoration with neural prosthetics," *Neuron*, vol. 52, (no. 1), pp. 205-20, Oct 5 2006.

[2] L.R. Hochberg, M.D. Serruya, G.M. Friehs, J.A. Mukand, M. Saleh, A.H. Caplan, A. Branner, D. Chen, R.D. Penn, and J.P. Donoghue, "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, (no. 7099), pp. 164-71, Jul 13 2006.

[3] M. Velliste, S. Perel, M.C. Spalding, A.S. Whitford, and A.B. Schwartz, "Cortical control of a prosthetic arm for self-feeding," *Nature*, vol. 453, (no. 7198), pp. 1098-101, Jun 19 2008.

[4] K. Patterson, P.J. Nestor, and T.T. Rogers, "Where do you know what you know? The representation of semantic knowledge in the human brain," *Nat Rev Neurosci*, vol. 8, (no. 12), pp. 976-87, Dec 2007.

[5] A. Martin, "The representation of object concepts in the brain," *Annu Rev Psychol*, vol. 58, pp. 25-45, 2007.

[6] S.V. Shinkareva, R.A. Mason, V.L. Malave, W. Wang, T.M. Mitchell, and M.A. Just, "Using FMRI brain activation to identify cognitive states associated with perception of tools and dwellings," *PLoS ONE*, vol. 3, (no. 1), pp. e1394, 2008.

[7] E.C. Leuthardt, G. Schalk, J.R. Wolpaw, J.G. Ojemann, and D.W. Moran, "A brain-computer interface using electrocorticographic signals in humans," *J Neural Eng*, vol. 1, (no. 2), pp. 63-71, Jun 2004.

[8] K.J. Miller, G. Schalk, E.E. Fetz, M. den Nijs, J.G. Ojemann, and R.P. Rao, "Cortical activity during motor execution, motor imagery, and imagery-based online feedback," *Proc Natl Acad Sci U S A*, vol. 107, (no. 9), pp. 4430-5, Mar 2 2010.

[9] G. Schalk, K.J. Miller, N.R. Anderson, J.A. Wilson, M.D. Smyth, J.G. Ojemann, D.W. Moran, J.R. Wolpaw, and E.C. Leuthardt, "Two-dimensional movement control using electrocorticographic signals in humans," *J Neural Eng*, vol. 5, (no. 1), pp. 75-84, Mar 2008.

[10] W. Wang, A.D. Degenhart, J.L. Collinger, R. Vinjamuri, G.P. Sudre, P.D. Adelson, D.L. Holder, E. Leuthardt, D.W. Moran, M.L. Boninger, A.B. Schwartz, D.J. Crammond, E.C. Tyler-Kabara, and D.J. Weber, "Human motor cortical activity recorded with micro-ECoG electrodes during individual finger movements," in Proc. IEEE EMBC, 2009, pp. Pages.

[11] K.J. Miller, S. Makeig, A.O. Hebb, R.P. Rao, M. denNijs, and J.G. Ojemann, "Cortical electrode localization from X-rays and simple mapping for electrocorticographic research: The "Location on Cortex" (LOC) package for MATLAB," *J Neurosci Methods*, vol. 162, (no. 1-2), pp. 303-8, May 15 2007.

[12] G. Schalk, D.J. McFarland, T. Hinterberger, N. Birbaumer, and J.R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Trans Biomed Eng*, vol. 51, (no. 6), pp. 1034-43, Jun 2004.

[13] D.J. McFarland, T. Lefkowicz, and J.R. Wolpaw, "Design and operation of an EEG-based brain-computer interface (BCI) with digital signal processing technology.," *Behav Res Meth Instrum Comput*, vol. 29, pp. 337-345, 1997.

[14] E. Edwards, S.S. Nagarajan, S.S. Dalal, R.T. Canolty, H.E. Kirsch, N.M. Barbaro, and R.T. Knight, "Spatiotemporal imaging of cortical activation during verb generation and picture naming," *Neuroimage*, Dec 21 2009.

[15] N.E. Crone, A. Sinai, and A. Korzeniewska, "High-frequency gamma oscillations and human brain mapping with electrocorticography," *Prog Brain Res*, vol. 159, pp. 275-95, 2006.

[16] D.A. Heldman, W. Wang, S.S. Chan, and D.W. Moran, "Local field potential spectral tuning in motor cortex during reaching," *IEEE Trans Neural Syst Rehabil Eng*, vol. 14, (no. 2), pp. 180-3, Jun 2006.

[17] T.M. Mitchell, *Machine Learning*: McGraw-Hill, 1997.

[18] B.J. Baars, "Language," in Cognition, Brain, and Consciousness, B. J. Baars and N. M. Gage eds.: Elsevier Ltd., 2007.

[19] D.M. Taylor, S.I. Tillery, and A.B. Schwartz, "Direct cortical control of 3D neuroprosthetic devices," *Science*, vol. 296, (no. 5574), pp. 1829-32, Jun 7 2002.

[20] W. Wang, S.S. Chan, D.A. Heldman, and D.W. Moran, "Motor cortical representation of position and velocity during reaching," *J Neurophysiol*, vol. 97, (no. 6), pp. 4258-70, Jun 2007.

[21] E. Donchin, K.M. Spencer, and R. Wijesinghe, "The mental prosthesis: assessing the speed of a P300-based brain-computer interface," *IEEE Trans Rehabil Eng*, vol. 8, (no. 2), pp. 174-9, Jun 2000.

[22] R.A. Andersen, J.W. Burdick, S. Musallam, B. Pesaran, and J.G. Cham, "Cognitive neural prosthetics," *Trends Cogn Sci*, vol. 8, (no. 11), pp. 486-93, Nov 2004.

[23] N.F. Ramsey, M.P. van de Heuvel, K.H. Kho, and F.S. Leijten, "Towards human BCI applications based on cognitive brain systems: an investigation of neural signals recorded from the dorsolateral prefrontal cortex," *IEEE Trans Neural Syst Rehabil Eng*, vol. 14, (no. 2), pp. 214-7, Jun 2006.

[24] K.J. Miller, "Broadband spectral change: evidence for a macroscale correlate of population firing rate?," *J Neurosci*, vol. 30, (no. 19), pp. 6477-9, May 12.