



## Decoding word and category-specific spatiotemporal representations from MEG and EEG

Alexander M. Chan<sup>a,b,\*</sup>, Eric Halgren<sup>c</sup>, Ksenija Marinkovic<sup>c</sup>, Sydney S. Cash<sup>b</sup>

<sup>a</sup> Medical Engineering and Medical Physics, Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA, USA

<sup>b</sup> Department of Neurology, Massachusetts General Hospital, Boston, MA, USA

<sup>c</sup> Department of Radiology, University of California San Diego, La Jolla, CA, USA

### ARTICLE INFO

#### Article history:

Received 28 July 2010

Revised 20 October 2010

Accepted 26 October 2010

Available online 30 October 2010

#### Keywords:

Language

Semantic category

Machine learning

Decoding

MEG

EEG

### ABSTRACT

The organization and localization of lexico-semantic information in the brain has been debated for many years. Specifically, lesion and imaging studies have attempted to map the brain areas representing living versus nonliving objects, however, results remain variable. This may be due, in part, to the fact that the univariate statistical mapping analyses used to detect these brain areas are typically insensitive to subtle, but widespread, effects. Decoding techniques, on the other hand, allow for a powerful multivariate analysis of multichannel neural data. In this study, we utilize machine-learning algorithms to first demonstrate that semantic category, as well as individual words, can be decoded from EEG and MEG recordings of subjects performing a language task. Mean accuracies of 76% (chance = 50%) and 83% (chance = 20%) were obtained for the decoding of living vs. nonliving category or individual words respectively. Furthermore, we utilize this decoding analysis to demonstrate that the representations of words and semantic category are highly distributed both spatially and temporally. In particular, bilateral anterior temporal, bilateral inferior frontal, and left inferior temporal-occipital sensors are most important for discrimination. Successful intersubject and intermodality decoding shows that semantic representations between stimulus modalities and individuals are reasonably consistent. These results suggest that both word and category-specific information are present in extracranially recorded neural activity and that these representations may be more distributed, both spatially and temporally, than previous studies suggest.

© 2010 Elsevier Inc. All rights reserved.

### Introduction

With the advent of functional neuroimaging techniques (e.g., PET and fMRI), numerous studies have been performed to investigate the neural basis of semantic representations. Neuroanatomical differences in the representation of specific semantic categories, especially living and nonliving objects, have been seen in both imaging and lesion studies (Caramazza and Mahon, 2003; Caramazza and Shelton, 1998; Chao et al., 1999; Dhond et al., 2007; Hauk et al., 2008; Martin and Chao, 2001; McCarthy, 1995; Shinkareva et al., 2008; Tranel et al., 1997; Warrington and McCarthy, 1983; Warrington and Shallice, 1984). Despite extensive work investigating the animate/inanimate distinction, the reported results are variable from study to study (Devlin et al., 2002; Moore and Price, 1999). Most studies agree that the left posterior middle temporal gyrus is activated in response to tools and man-made objects (Chao et al., 1999; Damasio et al., 1996; Martin et al., 1996; Moore and Price, 1999; Mummery et al., 1998, 1996; Perani et al., 1999), and that inferior temporal-occipital cortex

is activated for animals and natural stimuli (Chao et al., 1999; Damasio et al., 1996; Perani et al., 1995, 1999). However, results are conflicting with regard to the medial temporal surface, left medial frontal cortex, and parietal cortex; several studies suggest activation for animals in these areas (Damasio et al., 1996; Martin et al., 1996) while other studies find activation by man-made and nonliving objects (Chao and Martin, 2000; Mummery et al., 1998, 1996; Perani et al., 1995). Furthermore, many of the brain areas showing differential activation to living and nonliving stimuli are only reported in a single study.

The variability of previously reported results may be due, in part, to the statistical analysis of high-dimensional neuroimaging data. The traditional univariate statistical techniques used to analyze these data require correction for multiple comparisons to control for false positives, often making them insensitive to subtle, but widespread, effects within the brain. Therefore, univariate techniques may yield differing results depending on the specific responses elicited by the particular experiment performed. We hypothesized that a multivariate decoding analysis, which considers relationships between all features concurrently, would be able to detect distributed cortical areas that are differentially activated by living and nonliving objects.

In these previous studies, due to the constraints of the imaging modality, the temporal representation of these semantic categories

\* Corresponding author. Massachusetts General Hospital, Thier 423, 55 Fruit Street, Boston, MA 02114, USA.

E-mail address: [amchan@mit.edu](mailto:amchan@mit.edu) (A.M. Chan).

could not be investigated. Furthermore, fMRI and PET do not directly measure neural activity, but rather a metabolic correlate. Utilizing electroencephalography (EEG) and magnetoencephalography (MEG) allows for the study of both the spatial and temporal dynamics involved in the language processing. In this study, we recorded simultaneous EEG and MEG of healthy participants performing a language task to explore the differences in the neural representation of living and nonliving objects as well as individual words.

For successful decoding of multichannel EEG and MEG data, a classifier which is robust to high-dimensional data must be utilized. In this study, support vector machines (SVMs) were chosen to decode semantic category and individual word information from neural representations. SVMs are a family of nonlinear machine-learning algorithms that are commonly used to classify high-dimensional data sets (Vapnik, 1995). In combination with the multichannel electro/magneto-physiological recordings performed in this study, SVMs allow for a multivariate examination of the spatiotemporal dynamics of the processing of words and concepts. In this report, we use subject-specific decoders to study individual semantic representations, and subsequently examine the consistency between subjects and modalities using generalized SVM classifiers.

The successful decoding of semantic information from high-dimensional neural recordings not only allows for the study of language processing, but also has potential applications in the future development of language-based neuroprostheses. In this study, we further extend the SVM analysis by showing that a scalable “hierarchical tree” decoding framework, that sequentially decodes word properties to narrow the search space, improves on the single classifier decoding results, and may allow for the decoding of larger libraries of words and concepts.

## Materials and methods

### Participants and data collection

Nine right-handed, healthy male volunteers were recorded using simultaneous scalp EEG and MEG while performing auditory and visual versions of a language task. The two tasks were performed in two separate sessions, separated by an average of 4 months. Participants were native-English speakers between the ages of 22–30. This study was approved by the local institutional review board, and signed statements of consent were obtained from all subjects.

MEG was recorded using a 306-channel Elekta Neuromag Vectorview system (Stockholm, Sweden). Signals were digitized at 600 Hz and filtered from 0.1 to 200 Hz. Data from magnetometers and gradiometers were recorded, however only gradiometers were utilized in this study due to the lower noise in these sensors. Simultaneous EEG recordings were obtained from a 64-channel EEG cap at a sampling rate of 600 Hz with the same filter settings as the MEG recordings. EEG was recorded using a mastoid electrode reference but were converted to a bipolar montage to reduce noise.

### Language task

A visual (SV) and an auditory version (SA) of a language task were performed by each participant. A single trial involved presentation of a written word for 300 ms (in the SV task), or an auditory word 500 ms in length (in the SA task), followed by a fixation point. A shorter visual stimulus window was chosen in order to align the potentials related to lexico-semantic processing. Subjects were instructed to press a button if the presented word represented an object larger than 1 foot in any dimension (target trials; e.g., tiger, sofa), while refraining from responding to objects smaller than a foot (nontarget trials; e.g., cricket, lipstick). Exactly half of the trials involved words representing objects larger than 1 foot, requiring a motor response (target trials). This required subjects to access the

semantic representations of these particular words and potentially retrieve visuospatial or propositional knowledge of the associated object. Words were equally divided between living objects (animals and animal parts) and nonliving objects (man-made items). Half of the trials presented a novel word which was only shown once during the experiment while the other half of the trials presented 1 of 10 repeated words (each shown multiple times during the experiment).

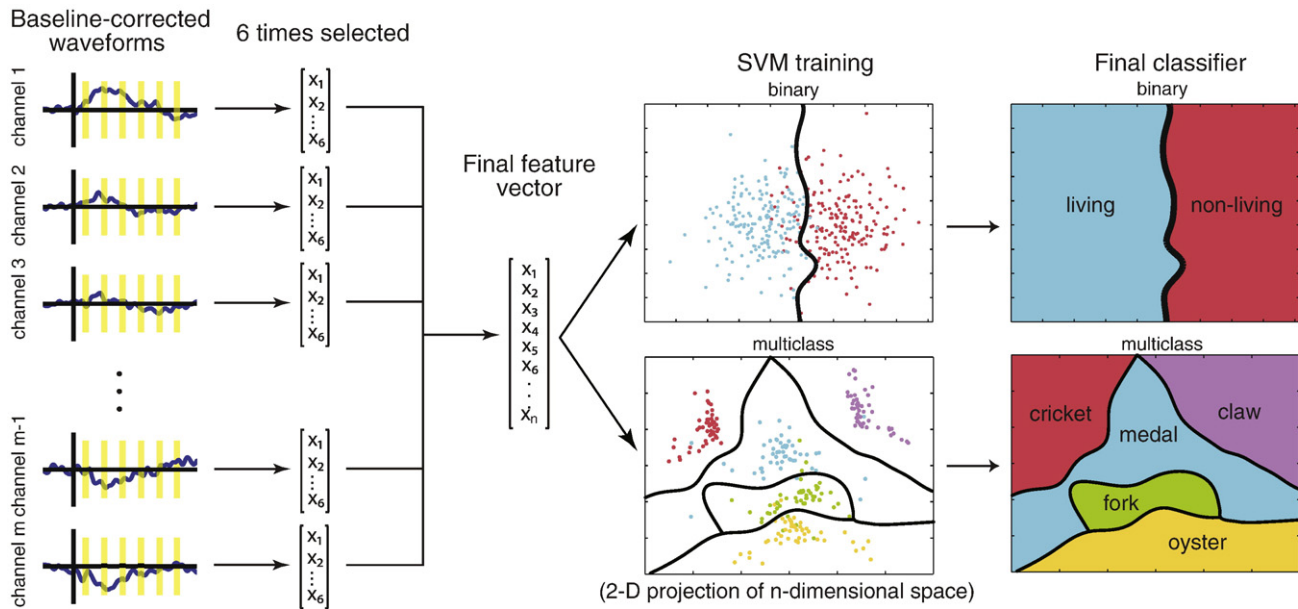
Novel words representing living and nonliving objects were balanced in terms of mean number of syllables (SA: living = 1.52, nonliving = 1.36, SV: living = 2.18, nonliving = 2.09), letters (SA: living = 5.22, nonliving = 5.21, SV: living = 6.49, nonliving = 6.8), and lexical frequency (SA: living = 15.5 per million, nonliving = 17.34, SV: living = 12.52, nonliving = 12.45) (Francis and Kucera, 1982). These word properties were not statistically different between living and nonliving object categories (Wilcoxon sign-rank,  $p > 0.05$ ). Auditory words had slightly fewer letters than visually presented words because they were required to fit within a 500 ms stimulus window. Repeated words were chosen to be representative of the novel words with respect to frequency and length. Visual stimuli were presented as white text on a black background while auditory stimuli were normalized in peak volume and length. The SV and SA tasks contained unique sets of words with no overlap between the two experiments. The visual version of the task included 390 trials while the auditory version included 780 trials. Analysis of modality-specific word processing in these tasks was previously performed by Marinkovic et al. (2003).

### Preprocessing

Signals from each channel of the MEG and EEG recordings were initially bandpass filtered from 1 to 30 Hz. Independent component analysis was performed on MEG and EEG signals, and EOG and EKG components were manually removed. For each trial, the continuous recordings were epoched from 1 s before to 2 s after stimulus onset. Trials containing large artifacts were rejected using a predefined amplitude threshold (300  $\mu$ V for EEG, 5pT/cm for MEG). Thresholds were intentionally set high to retain as much of the dataset as possible to reduce over-fitting the classifiers. After alignment to stimulus onset, waveforms from all channels were baseline corrected using a 500 ms prestimulus period. These preprocessing steps were performed within MATLAB, using the EEGLAB 6.03b (Delorme and Makeig, 2004) and FieldTrip toolboxes (ver. 20080611, <http://fieldtrip.fcdonders.nl/>).

### Decoding analysis

Two main components are necessary for the decoding of neural information: a feature extractor and a classifier. The goal of feature extraction is to reduce the full neural signal to a smaller number of components, or “features”, which are relevant for the subsequent classification task. In this study, the average amplitude in six 50-ms time windows were sampled from every channel and concatenated into a large feature vector for each trial (Fig. 1). Thus, a single feature vector represents the amplitude-based spatiotemporal properties of a single trial. The six time points selected for decoding living versus nonliving objects were 200, 300, 400, 500, 600, and 700 ms poststimulus, and the times selected for decoding individual words were 250, 300, 350, 400, 450, and 500 ms. Previous literature suggests that the N400 component of event related potentials (ERPs) is associated with semantic processing and integration, which informed our choice of these time ranges (Bentin et al., 1993; Hagoort et al., 2004; Kutas and Hillyard, 1984; Marinkovic et al., 2003). These time ranges also minimize early auditory or visual effects when examining individual repeated words and account for later activity when examining novel words (Marinkovic et al., 2003). Increasing the number of time points beyond six led to negligible



**Fig. 1.** Decoding framework utilizing amplitude-based feature extraction and SVMs. The amplitude at six poststimulus time points are selected from each channel and concatenated into an initial feature vector. The feature vectors from all channels are concatenated into a final feature vector. A single feature vector represents the spatio-temporal dynamics of a single trial. A nonlinear SVM is trained on these feature vectors to discriminate between the two semantic classes (living vs. nonliving objects) or between individual words. This results in a decision boundary by which new trials can be classified. In the multiclass case, multiple decision boundaries are generated to separate individual classes from each other.

increases in decoder performance, but substantially increased computational time needed to train the classifier. Because EEG and MEG signals have different magnitudes, features were normalized to take values between 0 to 1.

The second component of a decoding paradigm, the classifier, finds a relationship between the feature vector inputs and the corresponding word class (e.g., living/nonliving object category or specific word). The generated classifier model allows for the prediction of the word class for novel data. In this study, support vector machines (SVMs), as implemented by [Joachims \(1999\)](#), were chosen as the classifier. SVMs were chosen due to their robustness to high-dimensional data and ability to generate nonlinear decision boundaries ([Vapnik, 1995](#)). Using a set of training data from multiple classes (in our case, living and nonliving object categories or individual words), SVMs attempt to find a separating boundary which maximizes the margin between these classes ([Fig. 1](#)); this reduces over-fitting and allows for good generalization when classifying novel data.

To estimate the accuracy of the trained classifiers, a bootstrap cross-validation was performed. This procedure splits the data into nonoverlapping training and testing sets in order to evaluate the effectiveness of the classifier when encountering new data. For each round of cross-validation, 1–30 random trials of the same type (living or nonliving objects in the binary case, individual words in the multiclass case) were omitted before training the SVM. The omitted (test) trials were then individually classified using the resulting model and discriminant scores averaged to generate the final classification. One thousand rounds of bootstrap cross-validation were performed to obtain an estimate of classification accuracy. Only nonrepeated words were used to train the SVM to distinguish between living and nonliving objects, to allow for training on a large variety of unique stimuli within each category. By necessity, repeated words were used to train the SVM to classify individual words.

A radial-basis-function kernel, with parameter  $\gamma = 0.0005$ , was used in training the SVM to allow for nonlinear decision boundaries. The  $c$ -parameter, specifying the tradeoff between misclassification of training examples and maximal margin, was set to 1. A multiclass version of the classifier was also trained to discriminate between the five large (target) or five small (nontarget) words based on the implementation in [Crammer and Singer \(2002\)](#). Significance thresh-

olds, at  $p = 0.05$ , were computed using permutation distributions generated by performing 1000 repetitions of the cross-validation procedure on trials with shuffled target labels. All subsequent results indicating statistically significant decoding accuracies utilize this metric unless otherwise specified.

#### Data visualization

The final classifier generated by the SVM consists of a weight vector which can be used to classify new trials. In the linear case, the weight of each feature dictates the importance of that feature in the final classification. Thus, examining the linear SVM weights allows determination of important spatiotemporal features in the classification of living versus nonliving objects. By plotting the SVM weight vector on a 2-D topographic representation of the scalp (topoplot), we can generate a map of the time-sensor points which contribute the most to the final classifier.

In the individual word (multiclass) case, a single weight vector is generated for each of the five words. Highly variable weights associated with a particular feature indicate that the classifier more heavily utilizes this feature to discriminate different words. Thus, the variance of the weights for each feature was computed as a metric of relative importance of each time-sensor point and plotted in a topoplot. Because the lead field of a MEG planar gradiometer is directly under the sensors, these topoplots are a valid way of exploring the cortical areas contributing to the discrimination ([Hämäläinen and Ilmoniemi, 1994](#)).

Confusion matrices were also generated in the individual word (multiclass) case to compare the actual word presented with the word predicted by the classifier. These matrices indicate the type and quantity of errors generated by the multiclass SVM when decoding individual words and allow for a systematic analysis of the words that were most difficult to classify. Any given row of these matrices shows the distribution of classification of a particular word with the diagonal indicating correct classification and off-diagonals indicating errors.

#### Intermodality and intersubject decoding

To study supramodal representations, we trained an SVM on features from either the auditory or visual tasks from a single subject.



Data from the same subject, but opposing modality, was then used to test the classifier. Because no word overlap was present between the SV and SA versions of the task, this could only be done for the classification of categories (animals versus nonliving objects).

Intersubject decoding was also performed to examine the consistency of language-related representations between individuals. In this case, an SVM was trained on data from all but one subject within a single modality. The data from the remaining subject was then used as test data. This was repeated by omitting each subject in turn. This analysis was performed on both living/nonliving category and individual words.

#### *Hierarchical tree decoding*

Although utilizing a single multiclass decoder to distinguish individual word representations often works well, it does not directly incorporate a priori knowledge about semantic classes and the features which best discriminate these categories. To combine information from the classifier models generated to decode semantic category and individual words, we implemented a hierarchical tree framework which attempts to decode word properties sequentially. Given an unknown word, the tree decoder first classifies it as either a large (target) or small (nontarget) object. The word is then classified as living or nonliving object, and finally as an individual word within the predicted semantic category. This allows the appropriate features to be used to decode each word property, narrowing the search space before individual words are decoded. Furthermore, such a tree construct is easily scalable and could allow for the eventual decoding of large libraries of words.

As a proof-of-concept, a three-level hierarchical tree construct was implemented using a set of SVMs for each level of the tree. Amplitude (six time points from 200 to 700 ms) and spectral features (8–12 Hz power at six time points from 200 to 700 ms) were first utilized to decode whether an unknown word represented a large (target) or small (nontarget) object. The spectral features allowed for motor intent to contribute to this initial classification. In the second and third levels of the tree, amplitude features from 200 to 700 ms were utilized to decode living/nonliving object category, and amplitude features from 250–500 ms were used to decode individual words. A separate SVM was trained (using the same parameters described in section 2.4) for each branch of the tree using the appropriate trials. Upon classification of a new trial, the result of earlier levels determined which of the trained models would be utilized to decode subsequent features. For example, if a novel trial was first decoded as a large object (target trial), and subsequently decoded as a living object, the final classifier would label the trial as either a “dinosaur,” “python” or “steer”.

To compare performance to a single multiclass decoder, an SVM was trained to discriminate between all 10 words using the full set of amplitude and spectral features used in the hierarchical tree decoding. A bootstrap cross-validation with 1000 repetitions was again used to estimate the accuracy of this decoder.

## **Results**

### *Behavioral results*

To ensure that behavioral responses to different trial types did not contribute to the decoding of words and semantic categories, we first analyzed the accuracy and response times of button presses (to large objects) for all subjects. Accuracy of behavioral responses ranged from 71.6% to 95.5% with a mean of  $90.3 \pm 1.4\%$  across subjects. Mean response times varied from 760 to 1152 ms with a cross-subject mean of  $943 \pm 27$  ms. Mean accuracies for living and nonliving object categories across subjects were  $90.4 \pm 1.6\%$  and  $90.2 \pm 1.6\%$  respectively. Mean response times for living and nonliving object categories

were  $947 \pm 30$  ms and  $962 \pm 25$  ms. Accuracies and response times were not significantly different between living or nonliving object trials for any of the subjects (Wilcoxon sign-rank,  $p > 0.05$ ). It is therefore unlikely that differential behavioral responses influenced subsequent decoding analyses. Accuracies were not significantly different between SV and SA tasks (Wilcoxon,  $p > 0.05$ ), although mean response times were shorter for the visual task (SV: 864 ms, SA: 1023 ms, Wilcoxon,  $p < 0.00001$ ). As expected, response times were shorter for repeated versus novel words (repeated: 868 ms, novel: 1023 ms, Wilcoxon,  $p < 0.001$ ). Mean accuracies and times were not significantly different between individual repeated words for any subject (ANOVA,  $p > 0.05$ ).

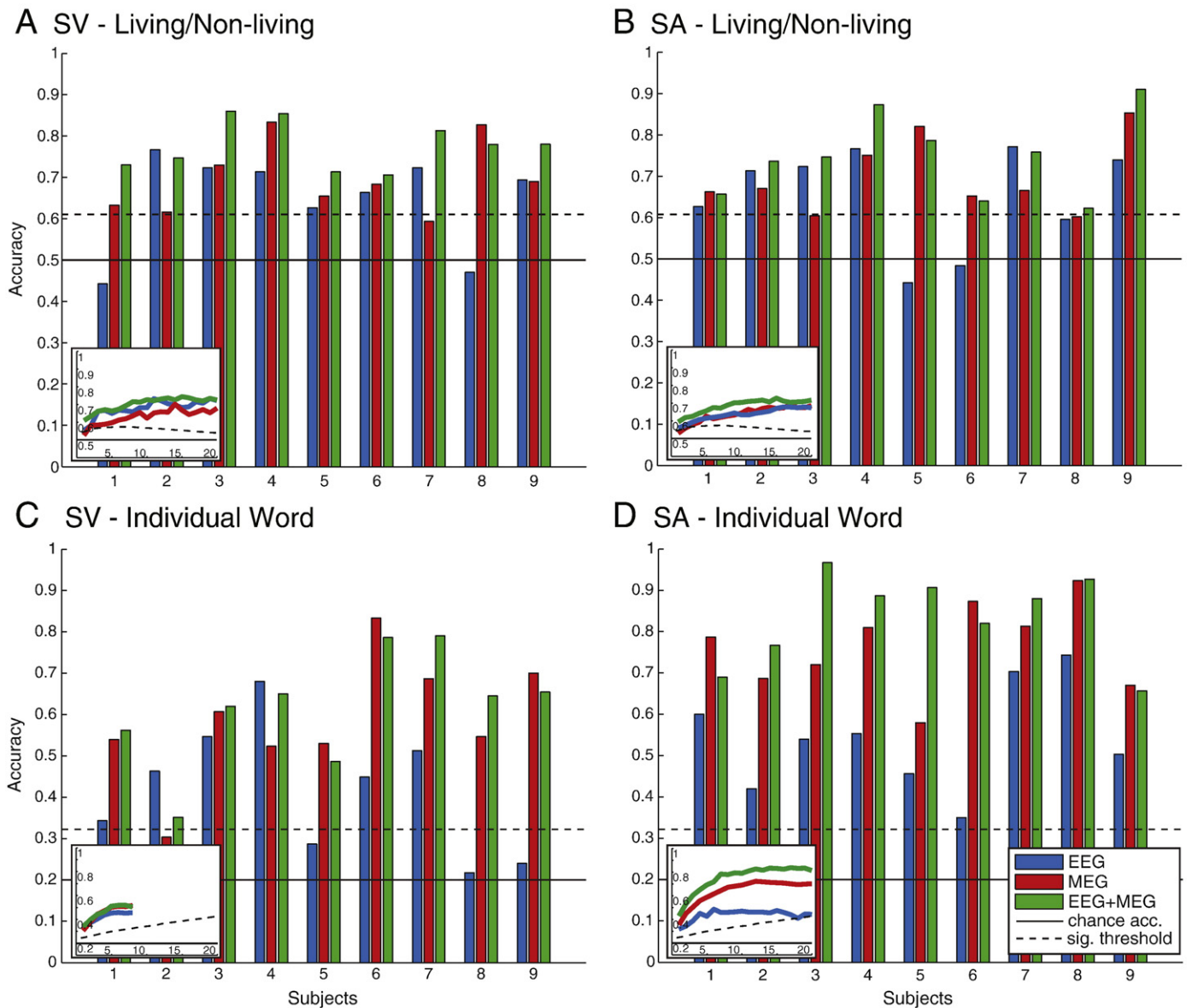
### *SVMs allow for decoding of semantic category*

We first attempted to train an SVM to decode living versus nonliving objects. The SVM was trained separately on EEG features, MEG features, and both combined. Fig. 2A–B illustrates the decode accuracies after averaging five trials (chance accuracy = 50%). When utilizing EEG features alone, data from seven of the nine subjects in the SV task and 6 of 9 in the SA task, showed statistically significant decoding accuracy (permutation test,  $p < 0.05$ ). When utilizing MEG features alone, data from eight of the nine subjects in SV and seven of nine in SA showed significant decoding accuracy (permutation test,  $p < 0.05$ ). Statistically significant decoding accuracy was obtained in data from all subjects when utilizing combined EEG and MEG features in both SV and SA tasks. When utilizing combined EEG and MEG features, accuracies ranged from 63% to 86% (mean  $\pm$  SE =  $76 \pm 2\%$ ) for the SV task and 62–91% (mean  $\pm$  SE =  $75 \pm 3\%$ ) for the SA task. Training on both MEG and EEG features increased accuracies by an average of 12% for the SV task and 10% for the SA task over using EEG features alone and 8% (SV) and 4% (SA) over MEG features alone (Wilcoxon sign-rank,  $p < 0.05$ ). Accuracies for the SV and SA task were not statistically different in any set of features when discriminating between living and nonliving objects (Wilcoxon,  $p > 0.05$ ). These results suggest that high-dimensional machine-learning algorithms, such as SVMs, are able to robustly extract semantic category information from multichannel electro/magneto-physiological recordings.

To explore the effect of the number of trials averaged on decoding accuracy, we also performed a leave-n-out cross-validation on all sets of features with all subjects (Fig. 2 inset panels). Not surprisingly, increasing the number of trials averaged resulted in increased decode performance in all cases. However, averaging more than approximately seven trials resulted in only marginal additional increases in performance.

### *SVMs accurately decode individual word representations*

We subsequently examined SVM decoding of individual word representations utilizing multiclass SVMs. We trained and tested classifiers on either the five repeated nontarget (small objects) or target words (large objects) to decode individual word representations without the potential motor confound (chance accuracy = 20%). The requirement for a motor action (button-press when the presented object was larger than one foot) may result in the decoding of that volitional response, rather than word processing information per se, when examining differences between all 10 words. The ability of the classifier to predict the observed word was statistically significant for all subjects after averaging five trials in at least one set of features (permutation test,  $p < 0.05$ ) (Fig. 2C–D). Accuracies varied from 32% to 79% (mean  $\pm$  SE =  $60 \pm 5\%$ ) using combined EEG/MEG features for the SV task (chance accuracy is 20%). For the auditory task, accuracies varied from 66% to 97% (mean  $\pm$  SE =  $83 \pm 4\%$ ). Training the SVM classifier on both EEG and MEG features increased average decode performance by 18% for the SV task and 29% for the SA task over using EEG features alone and 2% (SV) and 7% (SA) over MEG features alone (Wilcoxon,  $p < 0.05$ ).



**Fig. 2.** Decoding accuracy when distinguishing between living and nonliving objects or individual words. The bar graphs illustrate classifier accuracy for each subject when distinguishing between living and nonliving object category (A–B) or between individual words (C–D) after averaging five trials. Inset panels illustrate mean decoding accuracy as a function of the number of trials averaged. Blue indicates the use of EEG features, red indicates MEG features, and green indicates that both EEG and MEG features were used. In both the main figure and insets, chance accuracy (0.5 for living/nonliving and 0.2 for individual words) is shown as the horizontal black line and accuracies above the dashed line are statistically significant (permutation test,  $p < 0.05$ ). (A–B) Data from all subjects show significant decoding ability in at least one set of features. In all cases, utilizing combined EEG and MEG features resulted in significant decode accuracies. (C–D) When utilizing both EEG and MEG features, decoding performance when distinguishing individual words is statistically significant in all cases, and exceeds 95% accuracy in the SA task.

The decode accuracies of the SV and SA tasks when utilizing solely EEG features were not significantly different (Wilcoxon,  $p > 0.05$ ). However, utilizing MEG alone or both feature types resulted in significantly better performance in the SA data than utilizing the corresponding feature sets in the SV data (Wilcoxon,  $p < 0.01$ ).

The SA task contained twice as many trials as the SV task (780 for SA versus 390 for SV) which may have resulted in the difference in decoding accuracy between the two presentation modalities. By utilizing only the first 390 trials of the SA task, accuracy of the multiclass decoder after averaging five trials (mean  $\pm$  SE =  $61 \pm 4\%$ ) was not significantly different from SV performance (mean  $\pm$  SE =  $60 \pm 5\%$ ) (Wilcoxon,  $p > 0.05$ ).

Again, increasing the number of trials averaged increases decode performance substantially (Fig. 2C–D inset panels). In the case of individual word decoding for the SV task, there is a slight decrease in accuracy when the number of trials averaged is increased from 6 to 8.

This is likely due to the fact that increasing the number of trials averaged causes a corresponding decrease in the number of trials used for training the SVM, leading to a less robust classifier. This is especially pronounced in the multiclass SV case because of the relatively smaller number of total trials per condition when compared to the SA task. These data also illustrate that combining EEG and MEG features improves accuracy over either feature set alone. Taken together, these results demonstrate surprisingly robust ability to decode individual words from spatiotemporal features computed from multichannel electrophysiology.

*Linear probabilistic decoders are unable to handle high-dimensional data*

While a decoding analysis is a powerful method for exploring electro/magneto-physiological data, not all classification algorithms are suited for such an analysis. To demonstrate the advantages of

utilizing machine-learning techniques robust to high-dimensional data, we compared the decoding accuracy obtained when using SVMs (Sections 3.2 and 3.3) to the use of a popular probabilistic classifier. Because traditional Fisher linear discriminant analysis and Bayesian decoders are unable to handle cases in which the number of features is close to, or exceeds, the number of trials, we utilized a naïve Bayes classifier. Naïve Bayes classifiers assume independence of features, and are thus able to train and classify this particular set of MEG/EEG features.

When classifying living/nonliving category using MEG and EEG features, a naïve Bayes classifier resulted in average accuracies of  $54 \pm 4\%$  and  $51 \pm 3\%$  for SV and SA, respectively (chance = 50%). This was significantly lower than the SVM classification of the same data (76% for SV, 75% for SA, Wilcoxon sign-rank,  $p < 0.005$ ), and in fact not statistically different from chance. Similarly, when classifying individual words using MEG and EEG features, a naïve Bayes classifier yielded accuracies of  $41 \pm 4\%$  and  $46 \pm 3\%$  for SV and SA data respectively (chance = 20%). This, again, is significantly lower than the classification using an SVM (60% for SV, 83% for SA,  $p < 0.005$ ). These results suggest that a decoding analysis of MEG/EEG data requires techniques which are robust to high-dimensional data. In this case, SVMs, when compared to a naïve Bayes classifier, are better able to handle such data and can provide insight into the spatiotemporal representations of semantic knowledge.

#### *SVM weights show bilateral distributed cortical areas contribute to classification*

Examining the SVM weights allows us to determine the features which were most important in the generation of the final SVM classifier (Fig. 3). In the linear case, the weight of each feature dictates the importance of that feature in the final classification. Because the weights of a nonlinear classifier cannot be easily visualized, we utilized linear SVMs when examining classifier weights. The performance when using nonlinear SVMs was greater than the performance of the linear SVMs by 3.3% on average (Wilcoxon sign-rank,  $p < 0.05$ ), however decoding accuracy remained high in the linear case. In all cases where the nonlinear SVM yielded statistically significant decoding accuracy, the linear SVM also yielded statistically significant accuracy. Thus, examining the linear SVM weights allows determination of important spatiotemporal features in the classification.

Averaged weights across subjects for the visual (Fig. 3A) and auditory (Fig. 3B) tasks show a broadly distributed pattern of information-specific activity. Large weights are seen at all sampled time points and across both hemispheres. In particular, bilateral anterior temporal and inferior frontal weights increase to inanimate objects relative to living objects from 400 to 600 ms. A concurrent increase of SVM weights in response to living over nonliving objects is present at left inferior temporal-occipital sensors from 400 to 700 ms. Interestingly, an early temporal-occipital increase in weights to nonliving objects is seen at an earlier latency of 200 ms. While left inferior temporal-occipital activation to animals has been previously observed, the earlier activation to nonliving objects has not been reported.

When decoding individual word representations, the multiclass SVM generates one set of weights for each class. For visualization purposes, the variance of the SVM weights across words for each time-sensor point was computed and displayed (Fig. 3C–D). Features with higher variances differ more across classes, generally making them more important in the final classification. These data also show fairly distributed set of time-sensor points which contribute to the decoding. The SV data showed inferior occipital increase in weight variance from 300 to 400 ms, and inferior temporal activation from 400 to 500 ms (Fig. 3C). The SA task showed increased weight variance in bilateral anterior temporal areas from 250 to 450 ms with increases in posterior sensors at 300 and 500 ms (Fig. 3D).

#### *Systematic errors in individual word decoding reveal semantic structure*

Confusion matrices were constructed to analyze errors generated when discriminating between all 10 repeated words (Fig. 4A). The actual stimulus words are present along the vertical axis while the words predicted by the classifier are present along the horizontal axis. The colors along any given row (*actual word*) indicate the proportion of trials of that word which were classified as each of the possible choices (*predicted words*) (i.e., the confusion rate). Therefore, if the classifier correctly classified the word “feather” in all cases, the first element in the row corresponding to “feather” would be 1 (i.e., “feather” was always classified as “feather”) with all other elements being 0 (i.e., “feather” was never classified as any other word). Therefore, the diagonal elements in the matrix display correctly classified trials.

Visual examination of confusion matrices confirms that decoding of the MEG auditory data yields the highest accuracy, followed by EEG auditory data, followed by data from the visual task. The confusion matrices of combined EEG and MEG data were virtually identical to the confusion matrices generated to MEG data alone (data not shown). A larger confusion rate is visually apparent within target (large object) or nontarget (small object) classes (upper left and lower right corners), compared to between the two classes (lower left and upper right). The required motor response associated with the target trials may be providing additional nonlanguage information allowing for a decreased error rate when decoding between all 10 repeated words (as discussed in Section 3.3). Despite this, the ability to decode individual words is seen within the large and small object groups; this provides additional evidence that word-specific information is present in the neural signals being classified.

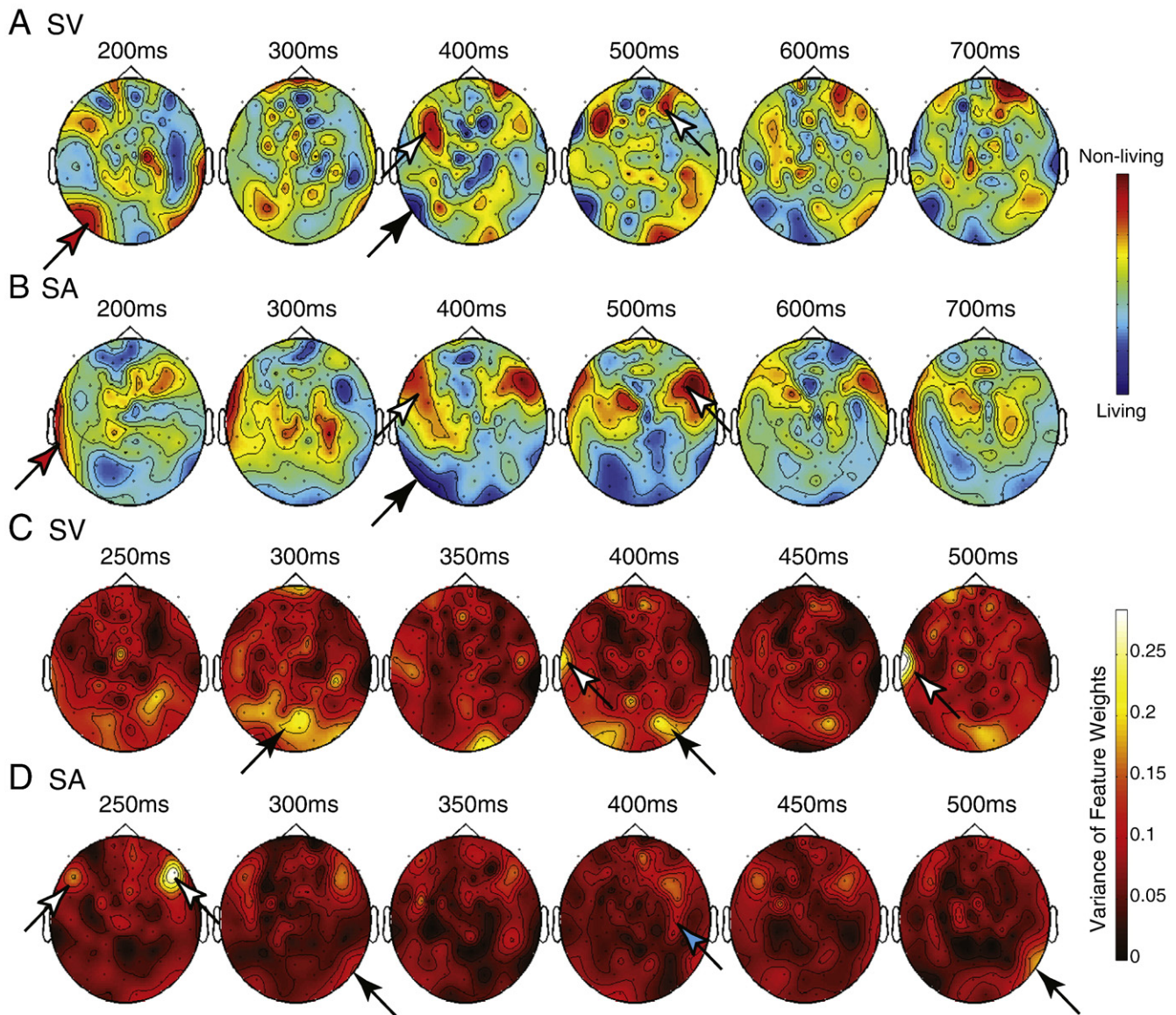
To quantify the effects of semantic category and large versus small objects on confusion rates, we performed a three-way ANOVA on these data (Fig. 4B). This was done to determine if two words which were within the same class (e.g., both living objects, both small objects, etc.) had a higher confusion rate than two words in different classes. In other words, the ANOVA compares differences in “within-class” confusion rates to “between-class” confusion rates. The ANOVA analysis involved three factors (living/nonliving, large/small, and subjects) with two levels in the categorical factors (within-class or between class) and nine levels in the subject factor (one for each subject).

For the SV task, the average large/small between-class confusion rate (mean  $\pm$  s.e. =  $0.0472 \pm 0.027$ ) was significantly smaller than large/small within-class confusion ( $0.125 \pm 0.045$ ;  $F = 45.72$ ,  $p < 0.00001$ ). Average living/nonliving object between-class confusion ( $0.074 \pm 0.037$ ) was significantly smaller than living/nonliving object within-class confusion ( $0.092 \pm 0.043$ ;  $F = 8.59$ ,  $p < 0.005$ ). For the SA task, the average large/small between-class confusion ( $0.038 \pm 0.028$ ) was significantly smaller than large/small within-class confusion ( $0.067 \pm 0.036$ ;  $F = 20.28$ ,  $p < 0.00001$ ). Average living/nonliving object between-class confusion ( $0.045 \pm 0.031$ ) was also significantly smaller than living/nonliving object within-class confusion ( $0.058 \pm 0.034$ ;  $F = 7.99$ ,  $p < 0.05$ ). This shows that it is more difficult for the classifier to discriminate words within the same semantic category than words of different categories. This suggests semantically related words have similar neural representations, and provides further evidence of the natural distinction between living and nonliving objects.

#### *Decoding is not based on low-level stimulus properties*

It is possible that the generated classifiers are utilizing neural activity related to low-level visual or auditory stimulus properties when decoding individual words. For example, the classifier may be decoding brain activity which is specific for the number of letters in the visual word or the number of syllables in the acoustic word, and not the semantic information associated with the word. To test this, we performed a shuffling based on stimulus properties to evaluate this potential confounding factor. Within either the five target or nontarget





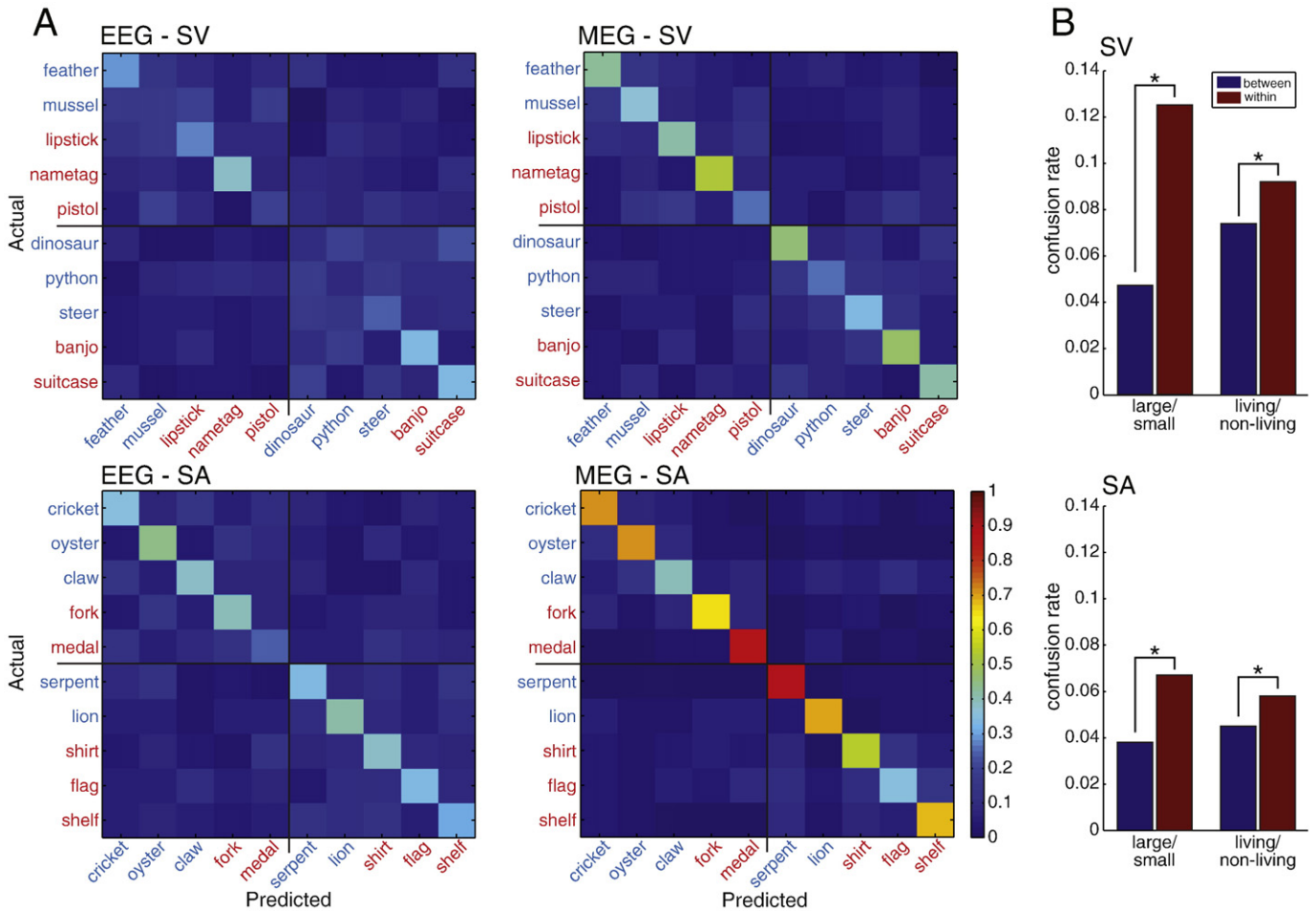
**Fig. 3.** Classifier weights show important times and locations for decoding. (A–B) SVM weights of the classifier trained on living and nonliving object categories in MEG show areas of significant differences. Areas of dark red indicate biases in classification towards nonliving objects, and blue denotes biases towards animals and living objects. Averaged weights across all subjects are shown at each sensor–time point for the (A) visual and (B) auditory tasks. Bilateral anterior temporal and inferior frontal differences are seen at 400–600 ms during both the SV and SA tasks (white arrows). Left temporal–occipital differences showing larger responses to objects are apparent at 200 ms (red arrows) with differences showing larger responses to living objects occurring at 400–700 ms in both modalities (black arrows). (C–D) Variance of SVM weights is shown at each time–sensor point indicating relative importance of each feature in individual word discrimination. Features with larger variance indicate larger separation between the SVM weights in that particular dimension and correlate with increased discrimination ability. (C) Extracranial weights from the SV task indicate occipital significance around 300–400 ms (black arrows) and inferior temporal significance at several times (white arrows). (D) Weights from the SA task show bilateral anterior temporal and inferior frontal significance from 250 to 450 ms (white arrows), and inferior occipital significance at 300 and 500 ms (black arrow). Inferior parietal significance is also seen from 350 to 400 ms (blue arrow).

words, we randomly swapped half of the trials between two words with equal numbers of letters or numbers of syllables, thus creating two categories with consistent sensory characteristics but scrambled lexical referents, while leaving the remaining three words unchanged. If the decode ability was solely based on either of these visual or phonetic properties of the stimulus, we would see no change in accuracy. In fact, the decoding accuracy of these sensory based categories dropped by 24% (letters) and 30% (syllables) (Wilcoxon sign-rank,  $p < 0.01$ ). Accuracies remained statistically above chance due to the fact that trials associated with three of the five words were left unchanged.

Although these low-level properties were not solely responsible for the decode ability, if these stimulus characteristics contributed information to the decoding, shuffling trials between two words with different sensory characteristics would result in a larger drop in

accuracy compared to shuffling between words with consistent sensory characteristics. The drop in performance when swapping trials between words with similar sensory characteristics was not significantly different from the performance when swapping trials between words with different sensory characteristics (25% for letters and 28% for syllables, Wilcoxon,  $p > 0.05$ ). This suggests that these sensory characteristics did not contribute significantly to the decoding of individual words in the visual version of the task.

We performed the same shuffling analysis for the SA task as well. The drop in performance was 23% when shuffling between words with the same number of syllables (Wilcoxon,  $p < 0.01$ ). This decrease in accuracy was not statistically different from the decrease in accuracy when shuffling between words with different numbers of syllables (20%, Wilcoxon,  $p > 0.05$ ).



**Fig. 4.** Individual word decoding confusion matrices. (A) Averaged confusion matrices for decoding all 10 individual words (averaging five trials) indicate the types of errors made. The vertical axis displays the actual stimulus word while the horizontal axis displays the word predicted by the classifier. The colors along any given row (*actual word*) indicate the proportion of trials of that word which were classified as each of the possible choices (*predicted words*). The diagonal elements display correctly classified trials. Words are sorted into small and large objects (divided by black lines), and living or nonliving categories (blue and red text). These matrices demonstrate a significant ability to decode individual words without regard to large/small conditions. (B) Within and between-category confusion rates are shown for the large/small and living/nonliving object distinctions. In all cases, confusion rates between categories are statistically lower than confusion rates within each category.

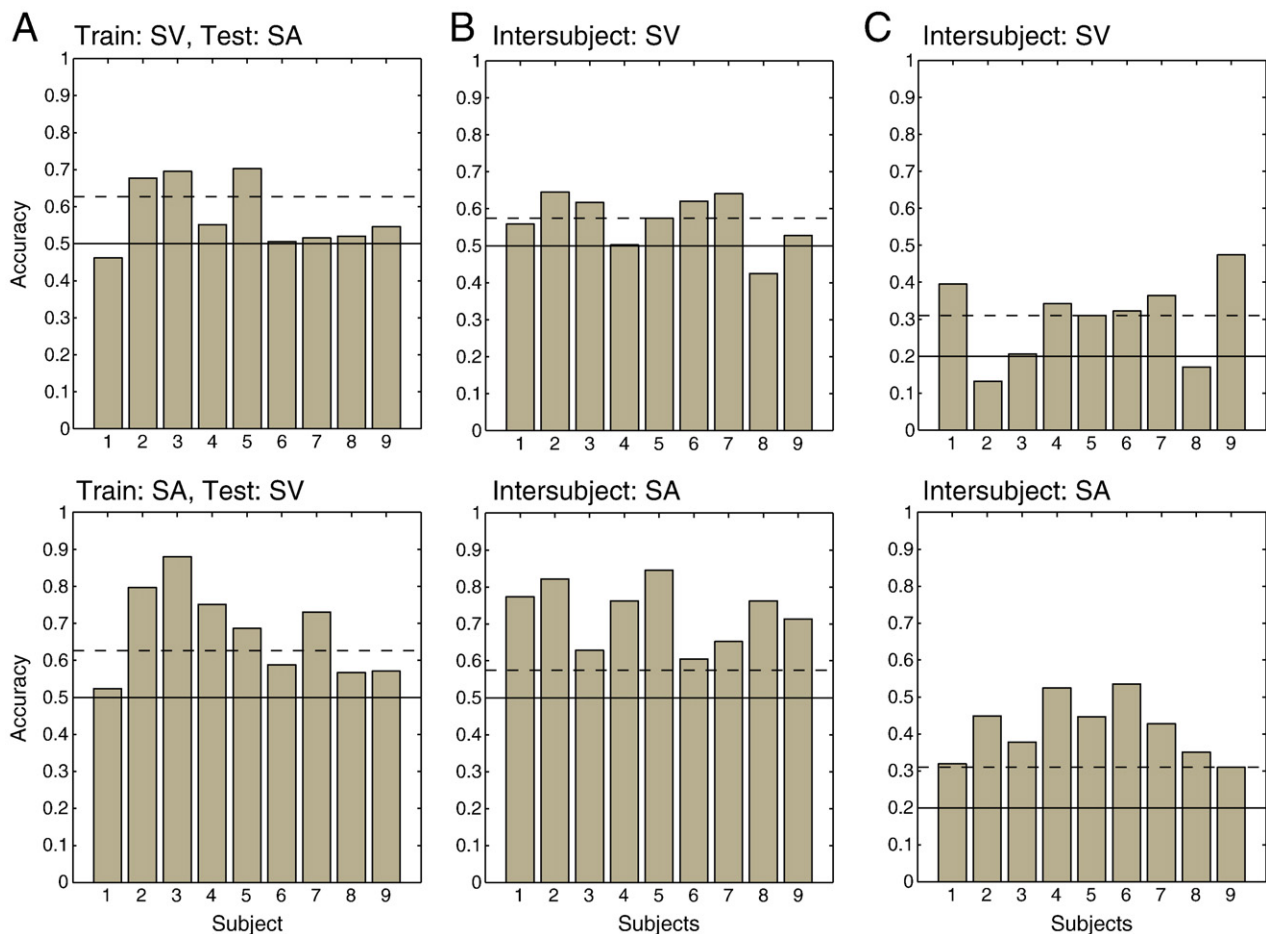
To control for the possibility of frequency-related acoustic properties of the words affecting the decode analysis (in the SA task), we attempted to predict stimulus properties using the same set of neural features used in the individual word decoding. In this case, the SVM algorithm performed a regression instead of classification to predict the power of the acoustic stimuli within five frequency bands (250–500 Hz, 500 Hz–1 kHz, 1–2 kHz, 2–4 kHz, and 4–8 kHz). If any of these acoustic properties contribute to the decoding of individual words, we would expect that an SVM trained on the previously used features would also be able to predict the power in these auditory frequency bands. To statistically test these results, a permutation distribution was computed by shuffling trials so that each trial was associated with a random set of stimulus band-power values for 2000 trainings of the SVM regression. The root mean square error was computed for each of these repetitions, resulting in a distribution of errors for the case that no information about stimulus band-power was present in the computed features. The root mean square error of this regression was not statistically significant based on a permutation distribution computed by shuffling the stimuli ( $p > 0.05$ , Supplementary Fig. S1). This result suggests that the decoding of individual words was not solely a result of differential representation of low-level properties of the auditory stimulus such as acoustic power.

#### *Inter-subject and intermodality decoding show shared neural representations of semantic information*

To investigate supramodal contributions to the generated classifiers, SVMs were trained on one stimulus modality and tested on the other modality. When training on visual data and testing on auditory data, statistically significant decode accuracies were obtained in three of nine subjects (Fig. 5A) with a mean accuracy across all subjects of  $57.5 \pm 3.0\%$ . When training on the auditory modality and testing on the visual modality, data from five of nine subjects showed significant decode accuracies with a mean accuracy across all subjects of  $67.7 \pm 4.1\%$ . This suggests that the models generated with features from either version of the task contain supramodal semantic information. This is more apparent in the case where the training set was larger and better able to produce a robust classifier (training on SA, testing on SV). By increasing the number of trials averaged, performance improves, as seen previously (Supplementary Fig. S2).

We also investigated the ability to train a generalized, subject-nonspecific decoder by training an SVM on data from all but one subject, and testing on the final subject's data. The accuracy obtained from such a cross-validation is an indication of the consistency of language-related representations between individuals. In the first case, an SVM was trained to discriminate between living and





**Fig. 5.** Intermodality and intersubject classification shows word and category representation consistencies. Accuracies of intermodality and intersubject decoding are shown after averaging 10 trials. Chance accuracy is indicated by solid horizontal line and the statistically significant threshold is shown by dashed line (permutation test,  $p < 0.05$ ). (A) Training on living/nonliving object data from SV and testing on data from SA results in data from 3 of 9 subjects showing statistically significant decode ability while training on SA and testing on SV results in data from five of nine subjects showing significant decode ability. This indicates supramodal semantic information is encoded within the classification models generated by the SVM. (B) Training an SVM on living/nonliving object data from all but one subject and testing on the final subject results in data from five of nine and nine of nine subjects showing statistically significant decode within the SV and SA modalities respectively. (C) Training an SVM on individual word representations from all but one subject and testing on the final subject results in data from six of nine and nine of nine subjects showing statistically significant decode ability. This indicates intersubject consistency in the neural representation of these words.

nonliving object categories. Data from five of nine subjects for SV and all subjects for SA showed statistically significant decoding performance (Fig. 5B,  $p < 0.05$ ). Mean accuracies were  $56.8 \pm 2.4\%$  and  $72.9 \pm 2.8\%$  for SV and SA, respectively.

A generalized SVM was also trained to discriminate between five large or small repeated words. Fig. 5C indicates that in six of nine cases for SV and all cases for SA, the decoding accuracy was significantly above chance levels. Mean accuracies were  $30.2 \pm 3.7\%$  for SV and  $41.3 \pm 2.7\%$  for SA (chance = 20%). Despite the fact that MEG sensor positions are variable between subjects, above-chance accuracies were obtained, suggesting that some word-specific information is consistent between individuals. Not surprisingly, however, subject-specific classifiers still yield significantly higher decode accuracies.

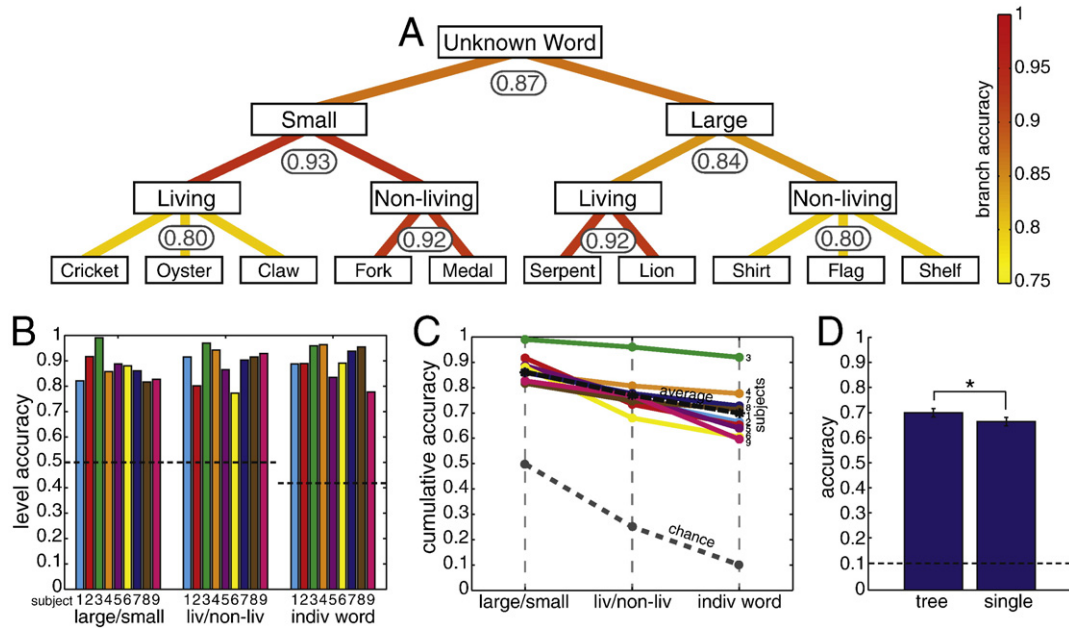
#### *Hierarchical tree decoding improves decoding performance*

To explore the potential practical use of machine-learning algorithms to decode larger libraries of words, we used SVM classifiers within the larger construct of a hierarchical tree decoder (Fig. 6). Such a paradigm is easily scalable and may allow for the eventual decoding of a large number of individual words or concepts. Utilizing a hierarchical tree decoding construct allows for the incorporation of a priori knowledge about semantic classes and the features which best discriminate these categories.

The average accuracy of all branches of the tree for the SA task was over 80% and accuracies at each level of the decoder were above 80% for all but two subjects (Fig. 6A–B). By examining cumulative accuracies at each level of the tree, we find that errors propagate from earlier levels, as expected, but accuracy ultimately remain above 60% in all cases (Fig. 6C). The mean overall accuracy of the tree decoder was 70%, significantly higher than the 67% accuracy of a single multiclass SVM trained on all 10 words (Wilcoxon sign-rank,  $p < 0.05$ ) (Fig. 6D). Data from all subjects, but subject 7, showed an improvement over the single SVM classifier when using the tree decoder. Thus, the hierarchical tree framework, by incorporating a priori knowledge of semantic properties, allows representations of individual word properties to be decoded more accurately than using a single multiclass decoder which treats each word as an independent entity.

#### **Discussion**

Understanding not only the spatial, but also the temporal representation of semantic categories and individual words requires analysis techniques robust to the high dimensionality of multichannel EEG and MEG data. In this study, we have demonstrated that a machine-learning technique, such as SVMs, can detect distributed differences in neural activity and robustly extract language-related



**Fig. 6.** Hierarchical tree decoding improves classification performance. A three-level hierarchical tree decoder was utilized to first decode the large/small distinction (utilizing amplitude and spectral features), then the living/nonliving object category (utilizing 200–700 ms amplitude features), and finally the individual word (utilizing 250–500 ms amplitude features). Data from decoding of the SA task are shown (data from SV task shown in [Supplementary Fig. S3](#)). (A) Average accuracies at each branch of the tree are shown with corresponding colors. Accuracies remain above 80% for all branches. (B) Accuracies at each level of the decoder are shown on a per subject basis with dotted lines indicating chance accuracy. (C) Cumulative accuracies at each level decrease as errors propagate through levels of the tree, but remain above 60%. (D) Performance of the hierarchical tree is a significant improvement (Wilcoxon sign-rank,  $p < 0.05$ ) over training a single multiclass SVM to discriminate between all 10 words.

information from electrophysiological recordings. These representations are supramodal and are relatively consistent between individuals. Utilization of a scalable hierarchical tree construct allows us to decode various word properties sequentially and further improves decode performance. These decoding techniques not only provide insight into the neural basis of language, but also may eventually be used as the basis of a language prosthetic device.

#### *Decoding analyses provide insight into spatiotemporal representation of words and categories*

Previous imaging studies investigating the neural basis of living/nonliving object representations have yielded variable results. The most consistent findings have been activation of the posterior middle temporal gyrus to tools and man-made objects and activation of inferior temporo-occipital cortex to animals (Chao et al., 1999; Damasio et al., 1996; Martin et al., 1996; Moore and Price, 1999; Mummery et al., 1998, 1996; Perani et al., 1995, 1999). Despite this, inconsistencies exist in the literature. While only a few studies report entirely conflicting results, many of the brain areas identified as showing differential activation to living or nonliving objects have only been reported in a single study. Moreover, one other MEG study failed to find any statistically significant differences between the perception of natural and man-made objects (Low et al., 2003). This may be due, in part, to experimental design, but may also be due to the statistical mapping analysis used to analyze these neuroimaging data. These techniques must correct for multiple comparisons and thus are most sensitive to brain areas which demonstrate large differences in activation between conditions. Often, multiple comparison corrections are based on spatial clustering, thus biasing the results toward contiguous arrays of activated voxels.

The SVM weights from our recordings are generally consistent with these previous imaging results, however, several important differences exist. The MEG data suggest that the bilateral anterior temporal, bilateral inferior frontal, and left parietal regions contribute to nonliving object category representation from 400 to 500 ms.

While bilateral temporal activations have been seen previously, frontal activity sensitive to object category have been largely seen only within the left hemisphere. Consistent with previous results, left inferior temporal-occipital SVM weights specific for living objects are apparent from 400 to 500 ms, but early 200 ms nonliving object-specific weights are also present in the same area. Activation to inanimate objects has not previously been seen in left inferior temporal-occipital cortex via functional neuroimaging. This finding suggests that a single brain area may respond to living and nonliving categories at different latencies, but more focal intracranial recordings may be necessary for further substantiating this hypothesis. Utilizing a recording modality with sufficient time resolution, and an analysis technique designed to handle high-dimensional data, allows for the discrimination of such time-separated effects. In contrast, the temporal blurring in fMRI and PET may only allow detection of the larger or more prolonged effect. This may explain discrepancies in previous imaging results; it is possible that the particular cognitive demands of each experimental task may elicit varying latencies of activity that manifest themselves differently in low time-resolution neuroimaging data.

Furthermore, while the N400 event-related potential (ERP) is known to be modulated by various semantic effects (Holcomb and Neville, 1990; Kutas and Hillyard, 1980, 1984), our results suggest that earlier components (possibly as early as 200–300 ms) may also contribute to the encoding of object category. This is especially pronounced in left inferior temporal-occipital sensors at 200 ms when classifying living versus nonliving objects in both visual and auditory modalities.

The results presented here also suggest a potential structure to the underlying representation of individual words. Because extracranial electrodes record the activity of large networks of concurrently active neurons, it is possible that the word-specific responses seen in our data are the superposition of many specific neural responses to lexico-semantic features of each word, as others have suggested (Caramazza et al., 1990; Pulvermüller, 2005; Tyler and Moss, 2001; Tyler et al., 2000). For example, the neural response to the word “banjo” may be

comprised of the sum of the specific activations related to how a banjo sounds, the visual characteristics of a banjo, the fact that “banjo” is a noun, and all the associative elements specific to the individual. One might expect that concepts with large amounts of overlapping characteristics would incorporate similar neural networks and thus would have similar macro-scale representations. The confusion matrix analysis supports this idea by indicating that, while the classifier was able to decode individual words, fewer errors were made between living and nonliving objects than within either of these semantic categories. This supports the intuitive notion that the representation of semantically related words may be more similar than the representation of words which are distant in semantic space. Under this hypothesis, it is not surprising that the novel trials separate nicely into living and nonliving object categories.

#### *Machine-learning techniques can robustly decode language information from extracranial neurophysiology*

This study also demonstrates the ability to extract semantic category and individual word representations from spatiotemporal features generated from noninvasive neurophysiology. The successful inter-modal classification shows that our machine-learning models are able to extract semantic information that is not specific to a single sensory modality. Not surprisingly, the cross-modality decode performance was lower than single modality performance; this may be partially due to differences in sensor placement and cognitive state between performance of the SV and SA tasks, often separated by months.

Despite the variability in recording conditions, intersubject classification also performed significantly above chance levels. This suggests that the representations of various objects, concepts, and semantic categories may be fairly consistent across individuals. However, it is important to note that the generalized decoder performs far worse than the subject-specific decoders. This may be due to variable electrode placement, variable cortical language-related representations, or both. This potential intersubject variability often decreases the sensitivity of traditional statistical mapping techniques used in imaging studies. Subject-specific decoding analyses, like the one presented in this study, overcome this by training and testing on a subjects own data.

A few other studies have attempted to decode word processing information from electrophysiology (Suppes and Han, 2000; Suppes et al., 1999, 1997). These studies utilize specifically chosen single channels of EEG. The minimum square error classifiers they used were therefore appropriate for decoding these low-dimensional data. In our case, however, feature vectors capture the entire spatiotemporal dynamics of each trial, and thus machine-learning techniques which are robust to high-dimensional data were necessary. We achieved higher average accuracies after averaging five trials than these previous studies have reported after averaging 10 trials (Suppes et al., 1997). In another study, Gonzalez Andino et al. (2007) also utilized SVMs to decode multichannel EEG recordings related to word and image processing. While the reported accuracies are impressive, the authors perform discrimination between distinct classes of stimuli (written words, pseudowords, line drawings, and scrambled images), rather than the more difficult task of decoding conceptual categories within a single stimulus modality.

While language information was extracted from both EEG and MEG recordings, MEG-based features yielded significantly higher accuracies. This differential accuracy of MEG versus EEG may simply be due to increased numbers of sensors in the MEG modality. However, despite the lower performance of the EEG features, combining EEG and MEG features improved performance over either recording modality alone, indicating that the information provided by EEG and MEG is not completely redundant. This suggests that neither recording modality is strictly superior to the other, and that EEG and MEG each provide unique information regarding neural processes.

This notion is supported by widespread evidence that MEG and EEG are sensitive to different neurophysiological processes (Cohen and Cuffin, 1983; Cuffin and Cohen, 1979; Deghani et al., 2010; Huang et al., 2007; Wolters et al., 2006).

While we have demonstrated that support vector machines are able to extract distributed language information from EEG and MEG recordings, not all multivariate classification techniques are equally successful. A naive Bayes classifier performs significantly poorer than the SVM, suggesting that analysis of EEG and MEG data requires algorithms which are robust against overfitting, and can handle high-dimensional data.

#### *Potential applications of language decoding*

The use of this decoding analysis not only provides insight into the nature of distributed language processing, but has implications for the development of a language-based neuroprosthesis. Machine learning algorithms, such as SVMs, can be trained on a patient's own data, making individual variability inconsequential. Furthermore, SVMs are robust to high-dimensional data, allowing for successful decoding broadly distributed semantic representations.

It is important to note, however, that this study is extremely preliminary with regard to the development of a practical communication prosthesis. Various practical barriers must be overcome before a language prosthesis is viable. The tasks used in this study are language comprehension tasks while a language prosthesis involves language production. However, in most models of language processing, the same underlying semantic representations of each of word are activated in both production and comprehension (Dell and O'Seaghdha, 1992; Indefrey and Levelt, 2004; Martin, 2003; Patterson et al., 2007). We have demonstrated that the representations we are decoding are supramodal, suggesting that semantic content is a major source of information in these recordings. This semantic representation is the desired decoding target for a language prosthetic device, so utilizing these language comprehension tasks as an initial pass in decoding analysis is not unreasonable. Furthermore, others have reported an ability to decode the motor commands associated with articulation, and we believe incorporating semantic information, as seen in this study, may greatly benefit such efforts (Guenther et al., 2009; Kellis et al., 2010).

An algorithm which narrows the search space of possible words by first determining various word properties (grammatical class, semantic category, visual attributes, etc.) before decoding individual concepts may require much less training and have higher performance. We have shown that this is possible with the hierarchical tree decoder and that performance improves as a result. The hierarchical framework presented here would allow for the decoding of a large library of concepts given the appropriate features to sequentially divide the search space. For example, concrete nouns and verbs produce different patterns of synchrony in EEG recordings (Weiss and Mueller, 2003), making coherence features a logical choice for discriminating this grammatical distinction. Given an adequate number of such distinctions, more realistically sized vocabularies may be utilized. The inclusion of a probabilistic syntactic/semantic language model, such as those used in automatic speech recognition (Baker, 1975), may further assist in narrowing the search-space and facilitate improved communication.

#### **Conclusion**

The decoding analyses used in this study allow for the study of distributed, but potentially subtle, representations of semantic information within the human cortex. These multivariate techniques offer advantages over traditional univariate statistical mapping analyses. We have shown that high-dimensional machine-learning techniques, in conjunction with EEG and MEG recordings, provide insight into both spatial and temporal aspects of language processing.



Furthermore, the ability to decode living/nonliving category or individual words between subjects and stimulus modality suggests that these representations are consistent and supramodal. We have also shown that utilizing word property information in an informed manner to decode individual words allows for increased performance and provides a potential framework for decoding larger libraries of words or concepts. These results, taken together, show that multivariate decoding techniques are a powerful tool for exploring distributed neural processing and the extraction of language information from electro/magneto-physiological recordings.

Supplementary materials related to this article can be found online at doi:10.1016/j.neuroimage.2010.10.073.

## Acknowledgments

This work was supported by an NDSEG Fellowship and a Frank H. Buck Scholarship to AMC and a Rappaport Fellowship to SSC. Overall support was provided by NIH grant NS18741. We thank J.M. Baker, A.R. Dykstra, C.J. Keller, F.E. McAllister, H. McAllister, N. Dehghani, J. Cormier, L.R. Hochberg, R. Zepeda, J. Donoghue, C. Sherman, C. Raclin, I. Sukhotinsky, S.S. Hou, and G.C. Sing for their helpful comments.

## References

- Baker, J., 1975. The DRAGON system—an overview. *IEEE Trans. Acoust. Speech Signal Process.* 23, 24–29.
- Bentin, S., Kutas, M., Hillyard, S.A., 1993. Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology* 30, 161–169.
- Caramazza, A., Mahon, B.Z., 2003. The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn. Sci.* 7, 354–361.
- Caramazza, A., Shelton, J.R., 1998. Domain-specific knowledge systems in the brain: the animate-inanimate distinction. *J. Cogn. Neurosci.* 10, 1–34.
- Caramazza, A., Hillis, A.E., Rapp, B.C., Romani, C., 1990. The multiple semantics hypothesis: multiple confusions? *Cogn. Neuropsychol.* 7, 161–189.
- Chao, L.L., Martin, A., 2000. Representation of manipulable man-made objects in the dorsal stream. *Neuroimage* 12, 478–484.
- Chao, L.L., Haxby, J.V., Martin, A., 1999. Attribute-based neural substrates in temporal cortex for lexical and knowing about objects. *Nat. Neurosci.* 2, 913–919.
- Cohen, D., Cuffin, B.N., 1983. Demonstration of useful differences between magnetoencephalogram and electroencephalogram. *Electroencephalogr. Clin. Neurophysiol.* 56, 38–51.
- Crammer, K., Singer, Y., 2002. On the algorithmic implementation of multiclass kernel-based vector machines. *J. Mach. Learn. Res.* 2, 265–292.
- Cuffin, B.N., Cohen, D., 1979. Comparison of the magnetoencephalogram and electroencephalogram. *Electroencephalogr. Clin. Neurophysiol.* 47, 132–146.
- Damasio, H., Grabowski, T.J., Tranel, D., Hichwa, R.D., Damasio, A.R., 1996. A neural basis for lexical retrieval. *Nature* 380, 499–505.
- Dehghani, N., Cash, S.S., Rossetti, A.O., Chen, C.C., Halgren, E., 2010. Magnetoencephalography demonstrates multiple asynchronous generators during human sleep spindles. *J. Neurophysiol.* 104, 179–188.
- Dell, G.S., O'Seaghdha, P.G., 1992. Stages of lexical access in language production. *Cognition* 42, 287–314.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Meth.* 134, 9–21.
- Devlin, J.T., Moore, C.J., Mummery, C.J., Gorno-Tempini, M.L., Phillips, J.A., Noppeney, U., Frackowiak, R.S., Friston, K.J., Price, C.J., 2002. Anatomic constraints on cognitive theories of category specificity. *Neuroimage* 15, 675–685.
- Dhond, R.P., Witzel, T., Dale, A.M., Halgren, E., 2007. Spatiotemporal cortical dynamics underlying abstract and concrete word reading. *Hum. Brain Mapp.* 28, 355–362.
- Francis, W.N., Kucera, H., 1982. *Frequency Analysis of English Usage: Lexicon and Grammar*. Houghton Mifflin, Boston.
- Gonzalez Andino, S.L., Grave de Peralta, R., Khateb, A., Pegna, A.J., Thut, G., Landis, T., 2007. A glimpse into your vision. *Hum. Brain Mapp.* 28, 614–624.
- Guenther, F.H., Brumberg, J.S., Wright, E.J., Nieto-Castanon, A., Tourville, J.A., Panko, M., Law, R., Siebert, S.A., Bartels, J.L., Andreasen, D.S., Ehirim, P., Mao, H., Kennedy, P.R., 2009. A wireless brain-machine interface for real-time speech synthesis. *PLoS ONE* 4, e8218.
- Hämäläinen, M., Ilmoniemi, R., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42.
- Hauk, O., Davis, M.H., Kherif, F., Pulvermuller, F., 2008. Imagery or meaning? Evidence for a semantic origin of category-specific brain activity in metabolic imaging. *Eur. J. Neurosci.* 27, 1856–1866.
- Holcomb, P.J., Neville, H.J., 1990. Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Lang. Cogn. Process.* 5, 281–312.
- Hagoort, P., Hald, L., Bastiaansen, M., Petersson, K.M., 2004. Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441.
- Huang, M.X., Song, T., Hagler Jr., D.J., Podgorny, I., Jousmaki, V., Cui, L., Gaa, K., Harrington, D.L., Dale, A.M., Lee, R.R., Elman, J., Halgren, E., 2007. A novel integrated MEG and EEG analysis method for dipolar sources. *Neuroimage* 37, 731–748.
- Indefrey, P., Levelt, W.J., 2004. The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.
- Joachims, T., 1999. Making Large-Scale SVM Learning Practical. In: Scholkopf, B., Burges, C.J.C., Smola, A.J. (Eds.), *Advances in Kernel Methods—Support Vector Learning*. MIT Press, Cambridge.
- Kellis, S., Miller, K., Thomson, K., Brown, R., House, P., Greger, B., 2010. Decoding spoken words using local field potentials recorded from the cortical surface. *J. Neural Eng.* 7, 056007.
- Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205.
- Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307, 161–163.
- Low, A., Bentin, S., Rockstroh, B., Silberman, Y., Gomolla, A., Cohen, R., Elbert, T., 2003. Semantic categorization in the human brain: spatiotemporal dynamics revealed by magnetoencephalography. *Psychol. Sci.* 14, 367–372.
- Marinkovic, K., Dhond, R.P., Dale, A.M., Glessner, M., Carr, V., Halgren, E., 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron* 38, 487–497.
- Martin, R.C., 2003. Language processing: functional organization and neuroanatomical basis. *Annu. Rev. Psychol.* 54, 55–89.
- Martin, A., Chao, L.L., 2001. Semantic memory and the brain: structure and processes. *Curr. Opin. Neurobiol.* 11, 194–201.
- Martin, A., Wiggs, C.L., Ungerleider, L.G., Haxby, J.V., 1996. Neural correlates of category-specific knowledge. *Nature* 379, 649–652.
- McCarthy, R.A., 1995. *Semantic Knowledge and Semantic Representations*. Psychology Press, New York.
- Moore, C.J., Price, C.J., 1999. A functional neuroimaging study of the variables that generate category-specific object processing differences. *Brain* 122 (Pt 5), 943–962.
- Mummery, C.J., Patterson, K., Hodges, J.R., Wise, R.J., 1996. Generating 'tiger' as an animal name or a word beginning with T: differences in brain activation. *Proc. Biol. Sci.* 263, 989–995.
- Mummery, C.J., Patterson, K., Hodges, J.R., Price, C.J., 1998. Functional neuroanatomy of the semantic system: divisible by what? *J. Cogn. Neurosci.* 10, 766–777.
- Patterson, K., Nestor, P.J., Rogers, T.T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987.
- Perani, D., Cappa, S.F., Bettinardi, V., Bressi, S., Gorno-Tempini, M., Matarrese, M., Fazio, F., 1995. Different neural systems for the recognition of animals and man-made tools. *NeuroReport* 6, 1637–1641.
- Perani, D., Schnur, T., Tettamanti, M., Gorno-Tempini, M., Cappa, S.F., Fazio, F., 1999. Word and picture matching: a PET study of semantic category effects. *Neuropsychologia* 37, 293–306.
- Pulvermuller, F., 2005. Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6, 576–582.
- Shinkareva, S.V., Mason, R.A., Malave, V.L., Wang, W., Mitchell, T.M., Just, M.A., 2008. Using fMRI brain activation to identify cognitive states associated with perception of tools and dwellings. *PLoS ONE* 3, e1394.
- Suppes, P., Han, B., 2000. Brain-wave representation of words by superposition of a few sine waves. *Proc. Natl Acad. Sci. USA* 97, 8738–8743.
- Suppes, P., Lu, Z.-L., Han, B., 1997. Brain wave recognition of words. *Proc. Natl Acad. Sci. USA* 94, 14965–14969.
- Suppes, P., Han, B., Epelboim, J., Lu, Z.-L., 1999. Invariance between subjects of brain wave representations of language. *Proc. Natl Acad. Sci. USA* 96, 12953–12958.
- Tranel, D., Damasio, H., Damasio, A.R., 1997. A neural basis for the retrieval of conceptual knowledge. *Neuropsychologia* 35, 1319–1327.
- Tyler, L.K., Moss, H.E., 2001. Towards a distributed account of conceptual knowledge. *Trends Cogn. Sci.* 5, 244–252.
- Tyler, L.K., Moss, H.E., Durrant-Peatfield, M.R., Levy, J.P., 2000. Conceptual structure and the structure of concepts: a distributed account of category-specific deficits. *Brain Lang.* 75, 195–231.
- Vapnik, V.N., 1995. *The nature of statistical learning theory*. Springer-Verlag, New York.
- Warrington, E.K., McCarthy, R., 1983. Category specific access dysphasia. *Brain* 106 (Pt 4), 859–878.
- Warrington, E.K., Shallice, T., 1984. Category specific semantic impairments. *Brain* 107 (Pt 3), 829–854.
- Weiss, S., Mueller, H.M., 2003. The contribution of EEG coherence to the investigation of language. *Brain Lang.* 85, 325–343.
- Wolters, C.H., Anwander, A., Tricoche, X., Weinstein, D., Koch, M.A., MacLeod, R.S., 2006. Influence of tissue conductivity anisotropy on EEG/MEG field and return current computation in a realistic head model: a simulation and visualization study using high-resolution finite element modeling. *Neuroimage* 30, 813–826.