

# EFFECTS OF AUDITORY ATTENTION ON LANGUAGE REPRESENTATION ACROSS THE HUMAN BRAIN

A DISSERTATION SUBMITTED TO  
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE  
OF BILKENT UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR  
THE DEGREE OF  
DOCTOR OF PHILOSOPHY  
IN  
ELECTRICAL AND ELECTRONICS ENGINEERING

By  
Özgür Yılmaz  
September 2019

EFFECTS OF AUDITORY ATTENTION ON LANGUAGE REPRESENTATION ACROSS THE HUMAN BRAIN

By Özgür Yılmaz

September 2019

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

---

Tolga Çukur(Advisor)

---

Emine Ülkü Sarıtaş Çukur

---

Emre Akbaş

---

Hüseyin Boyacı

---

Hüseyin Cem Bozşahin

Approved for the Graduate School of Engineering and Science:

---

Ezhan Karaşan  
Director of the Graduate School

# ABSTRACT

## EFFECTS OF AUDITORY ATTENTION ON LANGUAGE REPRESENTATION ACROSS THE HUMAN BRAIN

Özgür Yılmaz

Ph.D. in Electrical and Electronics Engineering

Advisor: Tolga Çukur

September 2019

Humans can effortlessly identify target auditory objects during natural listening and shift their focus between different targets. Unique allocation of brain resources would be inefficient for semantic search task. Here, we hypothesize that auditory attention shifts tuning of cortical voxels toward target category and that attention expands the representation of target words while compressing the representation of behaviorally irrelevant words across cortex. To test, we designed an fMRI experiment with a semantic search task. Subjects listened to natural stories twice while searching for words that are semantically related to either ‘humans’ or ‘places’. Fit voxelwise models for two attention tasks were compared to identify semantic tuning shifts in single voxels. Results indicate that attention shifts semantic tuning of single voxels broadly across cortex and attention warps language representation in favor of target words across cortex. We also introduced a novel feature regularization in voxelwise modeling for a naturalistic movie experiment. Feature regularization simply enforces similar model weights over semantically related stimulus features. We tested the proposed method on an fMRI experiment with naturalistic movies. Results suggest that the proposed method offer improved sensitivity in modeling of single voxels. Moreover, we proposed a novel method to improve the sensitivity of phase-sensitive fat-water separation in balanced steady-state free precession (bSSFP) acquisitions. In bSSFP applications using phased-array coils, reconstructed images suffer a lot from spatial sensitivity variations within individual coils. To improve, we first performed region-growing phase correction in individual coil images, then used a linear combination of phase-corrected images. Tests on SSFP angiograms of the thigh, lower leg, and foot suggest that the proposed method enhances fat–water separation in phased-array acquisitions with improved phase estimates.

*Keywords:* Computational Neuroscience, fMRI, Voxelwise Modeling, Tuning Shift, SSFP, Fat-Water Separation.

# ÖZET

## İŞİTSEL DİKKATİN DİLİN BEYİN ÜZERİNDEKİ GÖSTERİMİNE ETKİLERİ

Özgür Yılmaz

Elektrik ve Elektronik Mühendisliği, Doktora

Tez Danışmanı: Tolga Çukur

Eylül 2019

İnsanlar doğal dinleme sırasında kolaylıkla işitsel hedefleri saptayabilmekte ve farklı hedeflere odaklarını kaydırabilmektedir. Beyin kaynaklarının sabit bölüşümü durumunda işitsel arama performansı zorlaşır. Biz, hipotezimizde işitsel dikkatin kortikal voksellerin uyumlamasını hedef kategoriye kaydırıldığını ve hedef kelimelerin korteksteeki gösterimini genişlettiğini ileri sürmekteyiz. Hipotezimizi denemek için, anlamsal dikkat görevi içeren bir fMRI deneyi dizayn ettik. Denekler doğal hikayeleri bir kez ‘insanlar’ ile ilgili kelimelere bir kez de ‘mekanlar’ ile ilgili kelimelere dikkatlerini vererek dinledi. Sonuç olarak dikkatin korteks genelinde voksellerin anlamsal uyumlamasını kaydırıldığını ve dilin gösterimini hedef kelimeler lehine yamulttuğunu gözlemledik. Buna ilaveten vokal tabanlı modellemeye yeni bir öznelik düzenleme ekledik. Öznelik düzenleme anlambilimsel olarak benzer uyarı özneliklerinin benzer model ağırlıkları almasını zorlar. Sonuçlar bu yöntemin tekil voksellerin modellenmesinde gelişmiş hassasiyet sunduğunu gösterdi. Bunlara ilaveten, dengelenmiş kararlı-hal serbest devrim (bSSFP) kazanımlarında faza duyarlı yağ-su ayırımı geliştirecek yeni bir yöntem sunduk. Faz kaydırmalı bobin dizisi içeren bSSFP uygulamalarında, geriçatılmış görüntüler bireysel bobinlerdeki uzamsal hassasiyet değişkenliklerinden kötü etkilenmektedir. Yöntemimizde, önce ayrı ayrı bobin görüntülerinde bölge-büyütme faz düzeltme yaptık, sonra doğrusal düzeltilmiş görüntülerin optimal doğrusal kombinasyonunu elde ettik. Uyluk, alt bacak ve ayak SSFP anjiyogramlarında yaptığımız testlere göre önerdiğimiz yöntem faz-kaydırmalı kazanımlarda yağ-su ayırımı gelişmiş faz tahminleriyle iyileştirmektedir.

*Anahtar sözcükler:* Hesaplamalı Sinirbilim, fMRI, Voxel Tabanlı Modelleme, Uyumlama Kayması, SSFP, Yağ-Su Ayırımı.

## Acknowledgement

Firstly, I would like to express my sincere gratitude to my advisor Prof. Tolga ukur for the continuous support of my study, for his endless patience, motivation, and immense knowledge. His personal and academic guidance helped me in all the time of research and writing of this dissertation. I could not have imagined having a better advisor and mentor for my study.

Besides my advisor, I am thankful to the rest of my thesis committee, Prof. Emine lk Sarıtař ukur, and Prof. Tolga Esat zkurt for their invaluable comments and encouragement. I want to also thank my jury members, Prof. Hseyin Boyacı, Prof. Hseyin Cem Bozřahin, and Prof. Emre Akbař for all the insightful comments and suggestions they provide me to finalize my dissertation.

I would like to thank the following funding agencies for supporting the work in this dissertation: the Military Electronics Industries through ASELSAN PhD Scholarship, the Scientific and Technological Research Council of Turkey through TUBITAK BİDEB-2211, and the BAGEP Award of the Science Academy.

I want to thank my friends Kamil Nar, Sami Bolat, Alican Yksel, Toygan Kılı, Alper zaslan, Efe Ilıcak, Murat Kocaoęlu, Hseyin İnan, Ergin Din, Koray Ertan, mer Arol, Emin elik, and Salman Dar for all the memorable moments we lived together. I will always appreciate their support and genuine friendship. Special thanks to my dearest Kbra for her invaluable support throughout preparing my thesis presentation and my life in general.

I would like to thank my psychological counselors řule Atasoy and řafak Atay for their various contributions to my life. I am particularly grateful to Muhittin řiro for his unconditional support during my elementary education.

Last but not the least, I want to thank my beloved family: my mother Yıldız Yılmaz, my father Zeynel Yılmaz, my sister Gliz Yılmaz, and my uncle Sinan Yapku for their endless support throughout my entire life.

# Contents

|          |                                                                                           |          |
|----------|-------------------------------------------------------------------------------------------|----------|
| <b>1</b> | <b>Background</b>                                                                         | <b>1</b> |
| <b>2</b> | <b>Category-based Attention in Natural Speech Warps Semantic Maps in the Human Cortex</b> | <b>4</b> |
| 2.1      | Introduction . . . . .                                                                    | 4        |
| 2.2      | Methods . . . . .                                                                         | 8        |
| 2.2.1    | MRI Protocols . . . . .                                                                   | 8        |
| 2.2.2    | Subjects . . . . .                                                                        | 8        |
| 2.2.3    | Experimental Procedure . . . . .                                                          | 8        |
| 2.2.4    | ROI Abbreviations . . . . .                                                               | 9        |
| 2.2.5    | Functional Localizers . . . . .                                                           | 10       |
| 2.2.6    | Data Pre-processing . . . . .                                                             | 11       |
| 2.2.7    | Head Motion and Physiological Regressors . . . . .                                        | 12       |
| 2.2.8    | Semantic Model Construction . . . . .                                                     | 12       |
| 2.2.9    | Voxelwise Model Estimation and Validation . . . . .                                       | 13       |
| 2.2.10   | Construction of the Semantic Space . . . . .                                              | 14       |
| 2.2.11   | Tuning Shift Index . . . . .                                                              | 15       |
| 2.2.12   | Control Analysis in the Absence of Target Stimuli . . . . .                               | 16       |
| 2.2.13   | Behavioral Analysis . . . . .                                                             | 17       |
| 2.2.14   | Significance of Attentional Modulations . . . . .                                         | 17       |
| 2.2.15   | Construction of Control Models . . . . .                                                  | 18       |
| 2.2.16   | Cortical Flat Maps . . . . .                                                              | 18       |
| 2.3      | Results . . . . .                                                                         | 19       |
| 2.3.1    | Cortical Distribution of Attentional Modulations . . . . .                                | 22       |

|       |                                                                        |    |
|-------|------------------------------------------------------------------------|----|
| 2.3.2 | Attention Shifts Semantic Tuning towards the Target Category . . . . . | 23 |
| 2.3.3 | Attention Shifts Semantic Tuning in the Absence of Targets . . . . .   | 26 |
| 2.3.4 | Attention Warps Semantic Representation . . . . .                      | 26 |
| 2.3.5 | Discussion . . . . .                                                   | 27 |

|          |                                                                                                |           |
|----------|------------------------------------------------------------------------------------------------|-----------|
| <b>3</b> | <b>Informed Feature Regularization in Voxelwise Modeling for Naturalistic fMRI Experiments</b> | <b>38</b> |
| 3.1      | Introduction . . . . .                                                                         | 38        |
| 3.2      | Methods . . . . .                                                                              | 41        |
| 3.2.1    | MRI Protocols . . . . .                                                                        | 41        |
| 3.2.2    | Subjects . . . . .                                                                             | 42        |
| 3.2.3    | Experimental Procedure . . . . .                                                               | 42        |
| 3.2.4    | ROI Abbreviations . . . . .                                                                    | 42        |
| 3.2.5    | Functional Localizers . . . . .                                                                | 43        |
| 3.2.6    | Data Pre-processing . . . . .                                                                  | 44        |
| 3.2.7    | Cortical Flat Maps . . . . .                                                                   | 44        |
| 3.2.8    | Category Model . . . . .                                                                       | 45        |
| 3.2.9    | Voxelwise Model Estimation and Validation . . . . .                                            | 45        |
| 3.2.10   | Parameter Optimization . . . . .                                                               | 47        |
| 3.2.11   | JI-VM Model Estimation and Validation . . . . .                                                | 49        |
| 3.2.12   | Pseudo-code for JI-VM Implementation . . . . .                                                 | 50        |
| 3.2.13   | FI-VM Model Estimation and Validation . . . . .                                                | 52        |
| 3.2.14   | SI-VM Model Estimation and Validation . . . . .                                                | 52        |
| 3.2.15   | Pseudo-code for SI-VM Implementation . . . . .                                                 | 52        |
| 3.2.16   | Feature Coherence Analysis . . . . .                                                           | 53        |
| 3.2.17   | Spatial Coherence Analysis . . . . .                                                           | 54        |
| 3.2.18   | Noise Ceiling . . . . .                                                                        | 55        |
| 3.3      | Results . . . . .                                                                              | 55        |
| 3.3.1    | Prediction Scores of VM Methods . . . . .                                                      | 55        |
| 3.3.2    | Selectivity for Model Features . . . . .                                                       | 56        |
| 3.3.3    | Feature Coherence and Spatial Coherence of Model Weights . . . . .                             | 58        |
| 3.4      | Discussion . . . . .                                                                           | 61        |

|          |                                                                                                           |            |
|----------|-----------------------------------------------------------------------------------------------------------|------------|
| <b>4</b> | <b>Enhanced Phase-Sensitive SSFP Reconstruction for Fat-Water Separation in Phased-Array Acquisitions</b> | <b>70</b>  |
| 4.1      | Introduction . . . . .                                                                                    | 70         |
| 4.2      | Methods . . . . .                                                                                         | 72         |
| 4.2.1    | Dual-Acquisition PS-SSFP Reconstruction . . . . .                                                         | 72         |
| 4.2.2    | PS-SSFP for Phased-Array Acquisitions . . . . .                                                           | 73         |
| 4.2.3    | Alternative PS-SSFP Reconstructions . . . . .                                                             | 76         |
| 4.2.4    | Experiments . . . . .                                                                                     | 78         |
| 4.3      | Results . . . . .                                                                                         | 82         |
| 4.4      | Discussion . . . . .                                                                                      | 84         |
| <b>5</b> | <b>Conclusion</b>                                                                                         | <b>89</b>  |
| <b>A</b> | <b>Supplementary Figures-1</b>                                                                            | <b>107</b> |
| <b>B</b> | <b>Supplementary Figures-2</b>                                                                            | <b>122</b> |

# List of Figures

- 2.1 **Attention warps semantic representation of words in favor of attended categories.** The tuning-shift hypothesis predicts that attention warps semantic representation in favor target categories. During passive listening, words are represented in a continuous semantic space where semantically related words appear in nearby locations. Attention to a specific word category expands the representation of words that are semantically similar to the attended category, while compressing the representation of words semantically dissimilar to the attended category. . . . . 6
- 2.2 **Natural semantic search in spoken narratives.** (a) During the semantic search experiment, five subjects listened to three naturally spoken stories while attending to words that belonged to either one of two target categories: humans or places. Whole-brain BOLD responses were acquired. To estimate semantic tuning in single voxels, we fit voxelwise models that optimally predict measured BOLD responses in terms of the semantic features of the story stimuli. To account for hemodynamic delays, a finite impulse response filter was used for each semantic feature. The voxelwise tuning vectors were then averaged across different time delays. The resulting semantic tuning vectors describe how words in the story stimulus evoke BOLD responses. (b) Semantic tuning was estimated separately for each attention task. Tuning changes in single voxels are a unique, diagnostic aspect of the tuning-shift hypothesis. To test this hypothesis, we measured changes in voxel tuning between the two attention tasks. . . . . 7

2.3 **Attentional tuning changes for a representative voxel in intraparietal sulcus.** Semantic tuning vectors of a well-modeled voxel within IPS, in subject S1, for the two attention tasks. To simplify visualization, only a subset of words that are either neutral, semantically related to humans, or semantically related to places are shown. The color and location of the words reflect their coordinates in the semantic space, e.g., red-magenta words are related to the humans category whereas blue-purple words are related to the places category (see 2.2.10). The size of each word represents the magnitude of the evoked BOLD response. This voxel has a prediction score of 0.401 and exhibits significant attentional modulation ( $t$  test,  $q(\text{FDR}) < 0.05$ ). The voxel is strongly tuned for words related to the target category in both tasks, and weaker tuning is observed for behaviorally irrelevant words. . . . . 20

2.4 Caption next page. . . . . 31

2.4 **Attentional modulation of semantic tuning across cortex.**

(a) Distribution of semantic tuning during passive listening (subject S1, left hemisphere). The boundaries of functional ROIs are shown with white lines (See 2.2.4 for ROI abbreviations). Major anatomical landmarks and sulci are marked with gray lines. A three-dimensional semantic space was derived by principal component analysis on voxelwise tuning vectors measured during the passive listening task. Semantic tuning vectors were then projected onto the first three principal components, which were assigned to red, green and blue color channels. Yellow/green voxels are more selectively tuned for human-related words, whereas blue/purple voxels are more selectively tuned for place-related words. Regions of signal dropout, motor areas and insignificant voxels are shown in gray. (b) Distribution of semantic tuning during search for humans. While attending to humans, many voxels in posterior areas that are tuned to places under passive listening (e.g., Cingulate Sulcus, IPS) shift their tuning away from places. Voxels that are not tuned for humans under passive listening (e.g., in Precentral Sulcus, Frontal Gyrus, and FEF) shift their tuning toward humans. (c) Distribution of semantic tuning during search for places. While attending to places, many voxels in posterior areas that are tuned to humans under passive listening (e.g., Cingulate Sulcus) shift their tuning away from humans, and voxels that are not tuned for places under passive listening (e.g., in Frontal Gyrus, and Precentral Sulcus) shift their tuning toward places. These results indicate that many cortical voxels shift their semantic tuning toward target word categories. . . . . 32

2.5 Caption next page. . . . . 33

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |    |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.5 | <b>Attention shifts tuning of single voxels.</b> (a) To assess attentional changes in tuning of single voxels, we projected tuning vectors onto the first two dimensions of a semantic space. The semantic space was derived via principal component analysis (PCA) of voxelwise tuning vectors measured in a separate passive listening task. Only well-modeled voxels are shown (which exhibit significant attentional modulation and have prediction score greater than 1 s.d. above the mean). Each voxel is represented with a dot whose color indicates the tuning-shift index (TSI; see color bar in (b)). The projections of idealized human and place templates onto the semantic space are shown with two large dots. The marginal distributions of selectivity for humans and places are displayed with histograms in green. Many well-modeled voxels shift their tuning toward the target categories (Wilcoxon signed-rank test, $q(\text{FDR}) < 0.05$ ). (b) Voxelwise TSIs were projected onto the cortical flat map of the left hemisphere for subject S1. Voxels in many brain areas implicated in attention control including frontal operculum (FO), frontal eye fields (FEF), and intraparietal sulcus (IPS), in language comprehension including Broca’s area (Broca), Wernicke’s area (WER), supramarginal gyrus (SupMG), and inferior frontal gyrus (IFG) shift their tuning toward the target category. In contrast, voxels in posterior cingulate cortex (PCC) and angular gyrus (AngG) which partly overlap with the default mode network shift their tuning away from the target. This result suggests that the default mode network areas play a role in distractor suppression and error monitoring during semantic search tasks. (c) Histogram of the TSIs for well-modeled voxels (8401 voxels). Significantly positive TSIs are represented by dark gray bars and significantly negative TSIs are represented by light gray bars (Wilcoxon signed-rank test, $q(\text{FDR}) < 0.05$ ). The average TSI value is significantly greater than 0 (Wilcoxon signed-rank test, $P < 0.05$ ). These results indicate that attention shifts tuning of most well-modeled voxels toward the target. . . . . | 34 |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|

2.6 Caption next page. . . . . 35

2.6 **Tuning shifts in functional ROIs.** (a) Prediction scores are reported in functional ROIs as mean  $\pm$  s.e.m. across subjects. The average prediction score is  $18.65 \pm 1.75\%$  in language comprehension areas (Broca, SupMG, AngG, WER), and  $15.72 \pm 2.50\%$  in attention control areas (FO, FEF) (See 2.2.4 for ROI abbreviations). (b) Tuning shift indices are reported as mean  $\pm$  s.e.m. across subjects. Only significantly predicted voxels are considered ( $t$  test,  $q(\text{FDR}) < 0.05$ ) Attention causes varying levels of semantic tuning shifts in different functional ROIs. TSI is small in early auditory areas (AC), but larger in brain areas that correspond to later stages of language processing (WER and SupMG). The largest TSI values are seen in frontal areas that take part in language comprehension (Broca) and attention control (FO, FEF, IPS). This result implies that attentional tuning shifts become progressively stronger toward later stages of language processing and attention control. . . . . 36

## 2.7 Topology of the semantic space under passive listening and search tasks.

The tuning-shift hypothesis predicts that attention to a specific word category expands the semantic representation of target words while compressing the representation of the non-target words. To test this prediction, we estimated the topology of the semantic space independently under passive listening and the two attention tasks. PCA was performed separately on voxelwise semantic tuning vectors estimated during each task. The coefficients of a set of basic words in the first two PCs were used to visualize the topology of the semantic space. During attention tasks, words that are semantically related to the target category expand their representation, and words dissimilar to the target category are compressed. To quantify the degree of attentional warp in the semantic space, we computed a semantic warping index (see Methods). A warping index greater than 1 indicates that target words are more broadly spread in the semantic space compared to non-target words. A warping index less than 1 indicates that target words are more compactly spread in the semantic space compared to non-target words. We find that the semantic warping index is significantly greater than 1 consistently across subjects (Wilcoxon signed-rank test,  $P < 0.05$ ). These findings indicate that tuning shifts in cortex warp semantic representation in favor of attended words while compressing the representation of irrelevant words.

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |    |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.1 | <b>Natural movie experiment and model fitting.</b> (a) Whole-brain BOLD responses were acquired while subjects viewed natural movies. To estimate functional selectivity in single voxels, we fit voxelwise models that optimally predict measured BOLD responses in terms of the category features present in the movie stimuli [1]. The resulting models describe how each of 1705 object and action categories in the movie stimulus evoke BOLD responses. (b) Classical VM ignores potential correlations in stimulus features and correlations among neighboring voxels. FI-VM takes into account correlation among model features to increase accuracy in predicting single voxel selectivity. The resulting models have increased feature coherence. JI-VM further incorporates shared information between neighboring voxels. The resulting voxelwise models have both increased feature and spatial coherence. . . . .                                                                                                                                                                                                                                                                                                  | 40 |
| 3.2 | <b>Improvement in prediction performance over VM by FI-VM, SI-VM, and JI-VM.</b> The improvement in prediction performances for the three methods (FI- , SI- , JI-VM) over VM are shown for several well-known functional ROIs. Bar graphs display improvement in prediction scores ( $\Delta R^2$ , mean $\pm$ s.d. across subjects, see 3.2.9 for the definition of prediction scores). Each of the three methods significantly improves prediction scores in all ROIs ( $P < 0.05$ , Bootstrap test). For FI-VM, improvement in prediction score is higher in category-selective areas than in early visual areas ( $1.09 \pm 0.19$ vs $0.74 \pm 0.10$ ). For SI-VM, improvement is substantial in category-selective areas ( $2.75 \pm 0.55$ ) and in early visual areas ( $2.43 \pm 0.50$ ). JI-VM yields significantly higher prediction scores than all competing methods in all ROIs except V4 and RSC ( $P < 0.05$ , Bootstrap test). Improvement in prediction score is higher in category-selective areas than early visual areas ( $4.57 \pm 0.56$ vs $3.25 \pm 0.59$ ). These results imply that spatial and feature regularization improve model performance particularly across category-selective areas. . . . . | 57 |

3.3 **Cortical distribution of optimal L2-norm regularization parameter.** The  $L_2$ -regularization parameters over model features in four methods (VM, FI-VM, SI-VM, JI-VM) are visualized on the right hemisphere of subject S1. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by  $L_2$ -regularization. As a result, compared to VM, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for  $L_2$ -regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity. 59

3.4 Caption next page. . . . . 64

- 3.4 **Functional selectivity in a single voxel.** Functional selectivity of two well-modeled voxels in EBA are visualized for several object and action categories organized under the WordNet hierarchy. Each node represents the estimated response of the voxel to the labeled object or action category. Red nodes correspond to categories that evoke above-mean responses and blue nodes correspond to categories that evoke below-mean responses. Sizes of the nodes indicate the magnitude of the evoked responses. In the proposed approach, we incorporate feature and spatial regularization that limit unnecessary penalization of model weights by  $L_2$ -regularization ( $\lambda = 316.23$  vs.  $\lambda = 39.81$  for voxel-1 and  $\lambda = 2511.9$  vs  $\lambda = 39.81$  for voxel-2), while improving prediction scores ( $R^2 = 9.73$  vs.  $R^2 = 17.74$  for voxel-1 and  $R^2 = 2.12$  vs.  $R^2 = 11.01$  for voxel-2). As a result, in VM, the model weights are relatively suppressed compared to JI-VM. Furthermore, feature regularization enforces assignment of more similar weights to correlated categories in voxelwise models. For example, in voxel-1, model weights of the subordinate categories of ‘person.n.01’ (e.g., man, woman, child...) are closer to each other in JI-VM compared to VM. Taken together, these results imply increased sensitivity in capturing functional selectivity. . . . . 65
- 3.5 **Feature coherence for VM, FI-VM, SI-VM, and JI-VM.** Feature coherence for subject S1 is shown on the posterior part of the right hemisphere. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. Informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM is higher than that of VM and SI-VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITP. These results indicate that informed feature regularization better captures feature coherence broadly across visual cortex. . . . . 66

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |    |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.6 | <b>Feature coherence across functional ROIs.</b> Bar graphs display increases in feature coherence by three methods (SI-VM, FI-VM, JI-VM) over VM (mean $\pm$ std across subjects). Feature coherence of FI-VM and JI-VM models is significantly higher than that of VM in all functional ROIs ( $P < 0.05$ , Bootstrap test). The average increase in category-selective areas is $10.84 \pm 0.91\%$ for FI-VM and $17.29 \pm 1.42\%$ for JI-VM, whereas the average increase in early visual areas is $9.94 \pm 1.05\%$ for FI-VM and $13.67 \pm 1.13\%$ for JI-VM. In contrast, SI-VM that lacks feature regularization has relatively limited increase in feature coherence across all ROIs. These results suggest that informed feature regularization increases functional coherence in cortical representations particularly in high-level visual areas. . . . . | 67 |
| 3.7 | <b>Spatial coherence for VM, FI-VM, SI-VM and JI-VM.</b> Spatial coherence for subject S1 is shown on the posterior part of the right hemisphere. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, SI-VM and JI-VM yields higher spatial coherence consistently across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITP. These results suggest that spatial regularization better captures spatial coherence broadly across visual cortex. . . . .                                                                              | 68 |

3.8 **Spatial coherence across functional ROIs.** Bar graphs display increases in spatial coherence by three methods (FI-VM, SI-VM, JI-VM) over VM (mean  $\pm$  std across subjects). Spatial regularization enforces increased spatial coherence among cortical representations of neighboring voxels. SI-VM and JI-VM have significantly higher spatial coherence than VM in all functional ROIs ( $P < 0.05$ , Bootstrap test). The average increase in category-selective areas is  $19.78 \pm 1.45\%$  for SI-VM and  $9.37 \pm 1.40\%$  for JI-VM, whereas the average increase in early visual areas is  $21.99 \pm 1.41\%$  for SI-VM and  $8.06 \pm 1.75\%$  for JI-VM. Meanwhile, FI-VM that lacks spatial regularization does not yield any improvement in spatial coherence. These results indicate that spatial regularization increases spatial coherence among cortical representations of voxel neighborhoods. . . . . 69

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.1 | <b>Balanced Steady-State Free Precession.</b> (a) Transverse magnetization profiles of $\Delta\phi = \pi$ (dotted line) and $\Delta\phi = 0$ (dashed line) phase-cycled bSSFP sequences were simulated along with their complex-sum (CS, solid line). The simulations assumed $90^\circ$ flip angle, $TR/TE = 4.6/2.3$ msec and $T_1/T_2 = 1200/250$ msec at 1.5T. Magnitude (top row) and phase (bottom row) profiles are shown separately, and the locations of fat and water resonances are marked with arrows. Fat and water signals are $\pi$ -radians out of phase for $\Delta\phi = \pi$ but large field inhomogeneity may cause water or fat resonances to leak into neighboring bands. CS maintains the $\pi$ -radians fat-water phase difference while reducing sensitivity to field inhomogeneity. (b) To detect this phase difference, PS-SSFP uses a correction step to remove phase variations from additional sources including coil sensitivity. To illustrate the effects of coil sensitivity, in-phase bSSFP images ( $TR/TE = 9.2/4.6$ msec at 1.5T) were acquired using a quadrature extremity coil and an 8-channel phased-array coil (three representative channels are shown). Histograms of voxel magnitude (top row) and polar histograms of voxel phase (bottom row) are visualized. Because individual channels in the phased-array coil have compact spatial coverage, voxel magnitudes accumulate around lower intensities compared to the quadrature-coil image. The narrow sensitivity profiles also cause the voxel phases to be more broadly spread in phased-array images. | 75 |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|

- 4.2 **The proposed method.** (a) The initial step of the proposed reconstruction is to identify fat-voxel seeds near the peak of each coil's sensitivity profile. To do this, sensitivity maps are first estimated from the central portion of  $k$ -space data for each coil (bottom row). These maps are then multiplied with the corresponding bSSFP images (top row), and for each coil the voxel that maximizes the product is selected as the starting seed (marked with blue crosses). (b) Following coilwise seed selection, region-growing phase correction (RG-PC) is applied to individual-coil images ( $S_i$ ) in order to remove slowly varying phase components. The phase-corrected coil images ( $S_i^{pc}$ ) are then linearly combined, where images are weighted by the magnitude of respective coil sensitivities ( $|C_i|$ ). Phase-sensitive fat-water separation is finally performed based on the magnitude and phase of the combined image ( $S_{cmb}$ ). 77
- 4.3 **Dual-acquisition bSSFP images in the lower extremities.** The foot (a) and the lower leg (b), collected using an 8-channel phased-array coil. Unseparated complex-sum images are shown for bSSFP, whereas water images are shown for PS<sub>ms</sub> (multiseed, individual coil) and the proposed method. Note that PS<sub>ms</sub> was performed individually on each coil, and seed selection was identical across the two methods (starting seeds marked with blue crosses). In both a and b, the top row shows magnitude images and the middle row shows phase images (see colorbar) obtained from an individual coil. Meanwhile, the bottom row shows the final image combined across coils. Relatively large variations are observed in phase images away from the sensitivity peak of the coils: in posterior regions of the foot image and superior–anterior regions of the lower leg (marked with dashed ellipses). While these variations cause fat–water misclassification in individual-coil and combined PS<sub>ms</sub> images (marked with arrows), the proposed method produces more reliable phase estimates to improve tissue classification. . . 79

- 4.4 **In vivo thigh images combined across 8 channels of a phased-array coil.** First row: The unseparated bSSFP image shown as a reference. Second row: A coronal slice from fat images reconstructed using  $PS_{ss}$ ,  $PS_{ms}$ ,  $PS_{cc}$ , and the proposed method. Third row: A coronal slice from water images reconstructed using  $PS_{ss}$ ,  $PS_{ms}$ ,  $PS_{cc}$ , and the proposed method. Fourth row: Maximum-intensity projections (MIPs) across water images. Residual fat signals are seen in multiple regions of regular PS-SSFP reconstructions (marked with arrows). In contrast, the proposed method maintains reliable fat–water separation across the imaging volume. . . . . 81
- 4.5 **In vivo lower leg images combined across 8 channels of a phased-array coil.** First row: The unseparated bSSFP image shown as a reference. Second row: A coronal slice from water images reconstructed using  $PS_{ss}$ ,  $PS_{ms}$ ,  $PS_{cc}$ , and the proposed method. Third row: MIPs across water images. Fat–water misclassification is seen broadly across regular PS-SSFP reconstructions (marked with arrows). Meanwhile, the proposed method achieves improved fat–water separation across the entire imaging volume. 83
- 4.6 **Unseparated bSSFP images, fat/water images reconstructed using IDEAL, and fat/water images reconstructed using the proposed method.** First row: A sagittal slice from in vivo thigh images. Second row: A sagittal slice from in vivo lower leg images. Third row: A sagittal slice from in vivo foot images. . . . . 85

- A.1 Cortical flat maps of prediction score and TSI for subject S1.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S1. Prediction scores were calculated as Pearson's correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) are accurately predicted by the semantic model. (b) Cortical distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing. 108

**A.2 Cortical flat maps of prediction score and TSI for subject S2.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S2. Prediction scores were calculated as Pearson’s correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) were accurately predicted by the semantic model. (b) Cortical distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing. 109

### A.3 Cortical flat maps of prediction score and TSI for subject

**S3.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S3. Prediction scores were calculated as Pearson's correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) were accurately predicted by the semantic model. (b) distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing. . . . . 110

**A.4 Cortical flat maps of prediction score and TSI for subject S4.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S4. Prediction scores were calculated as Pearson’s correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) were accurately predicted by the semantic model. (b) Cortical distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing. 111

- A.5 Cortical flat maps of prediction score and TSI for subject S5.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S5. Prediction scores were calculated as Pearson's correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) were accurately predicted by the semantic model. (b) Cortical distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing. 112
- A.6 Prediction scores in functional ROIs.** Prediction scores are reported as mean  $\pm$  s.d. across subjects for the semantic model and the two control models, namely the spectrogram and phoneme models. In all functional ROIs, the average prediction score of the semantic model is significantly higher than that of the control models ( $P < 0.05$ , bootstrap test). Control models have relatively high prediction performances only in auditory cortex (AC). In high-level auditory and attention areas, the control models exhibit much poorer prediction performances than the semantic model. . . . 113

A.7 **Cortical flat maps of semantic tuning for subject S1.** Dis-  
tributions of semantic tuning are shown for passive listening and  
the two attention tasks. Each voxel’s tuning vector was projected  
onto the three semantic dimensions, and RGB colors are assigned  
to each voxel according to these projections. Insignificant voxels  
are shown in gray. Yellow/green voxels are tuned to humans while  
blue voxels are tuned to places and locations. Results indicate  
that many cortical voxels shift their semantic tuning toward the  
attended word category. . . . . 114

A.8 **Cortical flat maps of semantic tuning for subject S2.** Dis-  
tributions of semantic tuning are shown for passive listening and  
the two attention tasks. Each voxel’s tuning vector was projected  
onto the three semantic dimensions, and RGB colors are assigned  
to each voxel according to these projections. Insignificant voxels  
are shown in gray. Yellow/green voxels are tuned to humans while  
blue voxels are tuned to places and locations. Results indicate  
that many cortical voxels shift their semantic tuning toward the  
attended word category. . . . . 115

A.9 **Cortical flat maps of semantic tuning for subject S3.** Dis-  
tributions of semantic tuning are shown for passive listening and  
the two attention tasks. Each voxel’s tuning vector was projected  
onto the three semantic dimensions, and RGB colors are assigned  
to each voxel according to these projections. Insignificant voxels  
are shown in gray. Yellow/green voxels are tuned to humans while  
blue voxels are tuned to places and locations. Results indicate  
that many cortical voxels shift their semantic tuning toward the  
attended word category. . . . . 116

A.10 **Cortical flat maps of semantic tuning for subject S4.** Dis-  
tributions of semantic tuning are shown for passive listening and  
the two attention tasks. Each voxel’s tuning vector was projected  
onto the three semantic dimensions, and RGB colors are assigned  
to each voxel according to these projections. Insignificant voxels  
are shown in gray. Yellow/green voxels are tuned to humans while  
blue voxels are tuned to places and locations. Results indicate  
that many cortical voxels shift their semantic tuning toward the  
attended word category. . . . . 117

A.11 **Cortical flat maps of semantic tuning for subject S5.** Dis-  
tributions of semantic tuning are shown for passive listening and  
the two attention tasks. Each voxel’s tuning vector was projected  
onto the three semantic dimensions, and RGB colors are assigned  
to each voxel according to these projections. Insignificant voxels  
are shown in gray. Yellow/green voxels are tuned to humans while  
blue voxels are tuned to places and locations. Results indicate  
that many cortical voxels shift their semantic tuning toward the  
attended word category. . . . . 118

A.12 Caption next page. . . . . 119

A.12 **TSI histograms across cortical voxels.** (a) Main analysis. TSI distributions are shown for the cortical voxels (except motor areas) of subjects S1-S5. Only well-modeled (prediction score greater than 1 s.d. above the mean) and significantly attention-modulated ( $t$  test,  $q(FDR) < 0.05$ ) voxels are shown (7882-9014, depending on subject). Significantly positive TSIs (Wilcoxon signed-rank test,  $q(FDR) < 0.05$ ) are represented by dark gray bars and significantly negative TSIs (Wilcoxon signed-rank test,  $q(FDR) < 0.05$ ) are represented by light gray bars. Average TSI value is significantly greater than 0 in all subjects (Wilcoxon signed-rank test,  $P < 0.001$ ). These results indicate that attention shifts tuning of most well-modeled and attention-modulated cortical voxels toward the attended category. (b) Control analysis. To ensure that voxels shift their tuning towards attended categories even when targets are absent from the stimuli, a control analysis was done. Model fitting and TSI analyses were done using only those segments of the story stimuli that did not contain any words semantically related to humans or places. We then performed the TSI analyses on the fit models similarly to the main analysis. The TSI distributions in the absence of target stimuli are shown for the cortical voxels (except motor areas) of subjects S1-S5. Same voxel selection procedure is implemented as in (a). Average TSI value is significantly greater than 0 in all subjects (Wilcoxon signed-rank test,  $P < 0.001$ ). These results indicate that most well-modeled voxels shift their tuning toward the target category even when the target is not present in the stimuli. . . . . 120

A.13 **Distributions of tuning shifts across functional ROIs.** (a) Main analysis. Average TSIs of functional ROIs are shown for subjects S1-S5 (mean  $\pm$  s.e.m. across voxels). Results indicate that, attention causes different levels of semantic tuning shifts across functional ROIs. Areas that take place in attention control (FO, IPS, FEF), and language processing (Broca, WER) have high TSI for all subjects. (b) Control analysis. TSIs in functional ROIs are shown for subjects S1-S5 in the absence of targets in the stimuli (mean  $\pm$  s.e.m. across voxels). Results indicate that even when no target is present in the stimuli, functional ROIs have TSI characteristics similar to main analysis. . . . . 121

B.1 **Prediction scores of VM across functional ROIs.** The prediction performance of VM is shown for several well-known functional ROIs. Bar graphs display improvement in prediction scores ( $R^2$ , mean  $\pm$  std across subjects). VM has significantly positive prediction scores in all ROIs ( $P < 0.05$ , Wilcoxon signed rank test). The average prediction score is  $33.32 \pm 2.98$  in category-selective areas and  $10.37 \pm 1.54$  in early visual areas. . . . . 122

B.2 **Cortical distribution of optimal L2-norm regularization parameter for Subject S1.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity. . . . . 123

B.3 **Cortical distribution of optimal L2-norm regularization parameter for Subject S2.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity. . . . . 124

B.4 **Cortical distribution of optimal L2-norm regularization parameter for Subject S3.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity. . . . . 125

B.5 **Cortical distribution of optimal L2-norm regularization parameter for Subject S4.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity. . . . . 126

B.6 **Cortical distribution of optimal L2-norm regularization parameter for Subject S5.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity. . . . . 127

B.7 **Feature coherence in VM, FI-VM, and JI-VM for Subject S1.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex. . . . . 128

B.8 **Feature coherence in VM, FI-VM, and JI-VM for Subject S2.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITFP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex. . . . . 129

B.9 **Feature coherence in VM, FI-VM, and JI-VM for Subject S3.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITFP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex. . . . . 130

B.10 **Feature coherence in VM, FI-VM, and JI-VM for Subject S4.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITFP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex. . . . . 131

- B.11 Feature coherence in VM, FI-VM, and JI-VM for Subject S5.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex. . . . . 132
- B.12 Spatial coherence in VM, FI-VM, and JI-VM for Subject S1.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex. . . . . 133
- B.13 Spatial coherence in VM, FI-VM, and JI-VM for Subject S2.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex. . . . . 134

- B.14 Spatial coherence in VM, FI-VM, and JI-VM for Subject S3.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex. . . . . 135
- B.15 Spatial coherence in VM, FI-VM, and JI-VM for Subject S4.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex. . . . . 136
- B.16 Spatial coherence in VM, FI-VM, and JI-VM for Subject S5.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex. . . . . 137

# List of Tables

|     |                                                          |    |
|-----|----------------------------------------------------------|----|
| 4.1 | Quantitative Assessments of Fat-Water Separation . . . . | 84 |
|-----|----------------------------------------------------------|----|

# Chapter 1

## Background

Magnetic resonance imaging (MRI) is used to obtain images of biological tissues by means of strong magnetic fields. By modifying the pulse sequence of the scanner various tissue types can be distinguished such as blood vessels, bone, fat, gray and white matter in brain [2]. Apart from structural imaging, recently, functional neuroimaging studies are used to characterize the cortical activation associated to various mental tasks. Specifically, in functional Magnetic Resonance Imaging (fMRI) studies, temporal change in blood oxygenation is measured with a high spatial resolution. Blood oxygenation inside neural populations is strongly linked to the neural activities. Thus, scientists utilize fMRI to understand the functional characteristics of the human brain.

Many fMRI studies were conducted to localize specific cortical regions that are thought to represent low-level (retinotopic areas [3, 4], tonotopically organized auditory areas [5, 6, 7]) or high-level visual/auditory information (visual category areas [8, 9, 10, 11, 12, 13], auditory semantic areas [14, 15, 16]). In those studies, mostly, voxelwise modeling techniques were exploited to understand how stimulus features elicit Blood-oxygen-level-dependent (BOLD) responses across cortical voxels. Natural scene statistics or word statistics in text corpora were also used along with voxelwise modeling techniques to localize cortical regions that seem to represent visual or auditory information in case of naturally presented stimuli

[17, 18, 19, 20]. These studies indicated either specific cortical regions for each of various visual/auditory categories or a more continuous semantic space that represents stimulus information continuously across cortex. Specifically, in Huth et al. 2016, it is reported that the human brain represents thousands of word categories in a continuous semantic space mapped broadly across cerebral cortex.

In this thesis, chapter 2 mainly focuses on the effects of auditory attention on the language representation across cortex. For this purpose, we designed a functional magnetic resonance imaging (fMRI) experiment with a natural auditory semantic search task. Subjects listened to a diverse collection of naturally spoken stories. They were instructed to listen to each story twice while searching for words that are semantically related to either ‘humans’ (e.g., student, guy, parent, everyone) or ‘places’ (e.g., school, bar, north, lake). Whole-brain blood-oxygen-level-dependent (BOLD) responses were recorded via fMRI while subjects performed the search tasks. We used a powerful voxelwise semantic model that estimates BOLD responses in terms of the semantic features of the story stimuli. Later, we measured semantic tuning in single voxels of each subject by comparing the fit models under two attention conditions. Results indicate that auditory attention shifts semantic tuning toward the target category broadly across cortex. Another finding is that attention warps language representation across cortex in favor of target words, while compressing the representation of semantically irrelevant words.

Chapter 3 focuses on a novel voxelwise modeling approach for naturalistic fMRI experiments. In this approach, we introduced an informed feature regularization that takes into account correlations among model features. The informed feature regularization simply enforces similarity of model weights between features that are semantically related to each other for each voxel. In the proposed method, we further implemented an improved spatial regularization across voxelwise models that considers cortical distances among voxels instead of Euclidean. We evaluated performance of the proposed method on an fMRI dataset gathered from a naturalistic movie experiment. Results suggest that the proposed method offer improved sensitivity in modeling of single voxels in naturalistic fMRI experiments together with increased feature coherence and spatial coherence of voxelwise models.

Apart from these studies, in chapter 4, we propose a novel method to improve the sensitivity of phase-sensitive fat-water separation in balanced steady-state free precession (bSSFP) acquisitions. bSSFP sequences typically generate higher levels of signal from fat when compared to water tissues. Separation of these two signals is critical for many clinical applications [21, 22, 23, 24]. However, in applications using phased-array coils, reconstructed images suffer a lot from spatial sensitivity variations within individual coils. To improve these reconstructions, in our proposed method, we first perform region-growing phase correction in individual coil images via unsupervised selection of a fat-voxel seed near the peak of each coil’s sensitivity profile. We then use an optimal linear combination of phase-corrected images to segregate fat and water signals. The proposed method was demonstrated on noncontrast-enhanced SSFP angiograms of the thigh, lower leg, and foot. Results indicate that the proposed method enhances fat–water separation in phased-array acquisitions by producing improved phase estimates across the imaging volume. Note that this chapter is based on a paper published in 2015 [25].

## Chapter 2

# Category-based Attention in Natural Speech Warps Semantic Maps in the Human Cortex

### 2.1 Introduction

Humans can effortlessly identify target auditory objects during natural listening and shift their focus between different targets [26]. Auditory attention is thought to mediate this remarkable ability by enhancing representations of auditory information in the brain. Many distinct types of auditory information from low-level tonal features to high-level semantic features are thought to be represented hierarchically across a network of regions including primary auditory cortex, temporal cortex, parietal cortex and prefrontal cortex [5, 6, 7, 27, 28, 29, 30, 31]. Previous studies using low-level spatial or tonal attention tasks reported response baseline or gain changes in primary auditory cortex [32, 33, 34] and parietal cortex [35, 36, 37]. High-level attention tasks were also reported to cause similar baseline or gain changes in temporal gyrus, frontal gyrus, and prefrontal cortex [38, 39, 40]. While baseline or gain modulations can elevate target responses above background signals, they cannot fully account for the dynamic nature of

real-world search [41].

Theoretical and physiological accounts suggest attentional contribution from a dynamic matched filter mechanism that warps representations at the population level particularly towards later stages of processing [41, 42, 43, 44]. A matched filter can expand the representation of attended features while compressing the representation of unattended ones across neocortex [1]. In accordance with these accounts, recent neurophysiology [45, 46, 47] and neuroimaging [48, 49, 50, 51, 52] studies have reported shifts in spectro-temporal receptive fields within primary auditory cortex during low-level spatial and tonal attention tasks. However, little is known about whether high-level auditory attention operates through a similar matched filter mechanism to warp auditory semantic representations and how these representational warps are distributed across cortex.

Providing evidence for the matched-filter hypothesis, in previous studies from The Gallant Lab, the authors found that semantic information about thousands of visual object categories is represented in smooth gradients across cortex, and that category-based visual search in natural movies causes voxelwise tuning shifts that warp semantic representations in favor of attended categories [1, 53]. They also found that during passive listening of natural stories, the human brain represents thousands of word categories in a continuous semantic space mapped broadly across cerebral cortex [20]. Thus, it is likely that high-level auditory semantic search for word categories in natural stories also induces broad voxelwise tuning shifts that warp semantic representations in favor of attended categories (Fig. 2.1).

To examine modulations in auditory semantic representation with improved sensitivity, we designed a natural auditory semantic search task (Fig. 2.2). Subjects listened to a diverse collection of naturally spoken stories that contained 1463 distinct words. They were instructed to listen to each story twice while searching for words that are semantically related to either ‘humans’ (e.g., student, guy, parent, everyone) or ‘places’ (e.g., school, city, north, lake) in alternating order across scans (see 2.2.3). Whole-brain blood-oxygen-level-dependent (BOLD) responses were recorded via functional magnetic resonance imaging (fMRI) from

five human subjects while they performed the search tasks (see 2.2.3).

To measure semantic tuning in single voxels in each subject, we used a powerful voxelwise semantic model that estimates BOLD responses in terms of the semantic features of the story stimuli [54]. The semantic features were identified by constructing a word embedding space based on word co-occurrence statistics in a large corpus of text. Each word in the stories was projected onto this embedding space. Regularized linear regression was used to fit voxelwise models that optimally predict the measured BOLD responses in terms of the semantic features. Two separate semantic models were obtained based on data collected during search for humans and during search for places. Model weights obtained under different attention tasks were compared to identify semantic tuning shifts in single voxels.

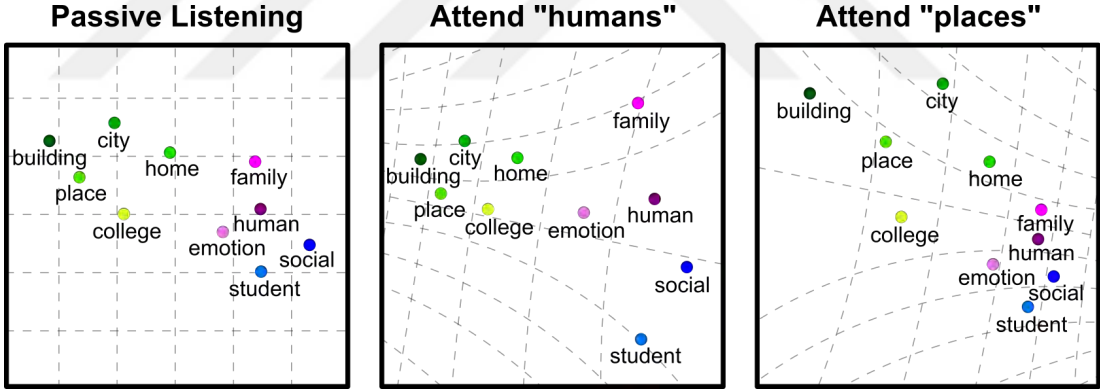


Figure 2.1: **Attention warps semantic representation of words in favor of attended categories.** The tuning-shift hypothesis predicts that attention warps semantic representation in favor target categories. During passive listening, words are represented in a continuous semantic space where semantically related words appear in nearby locations. Attention to a specific word category expands the representation of words that are semantically similar to the attended category, while compressing the representation of words semantically dissimilar to the attended category.

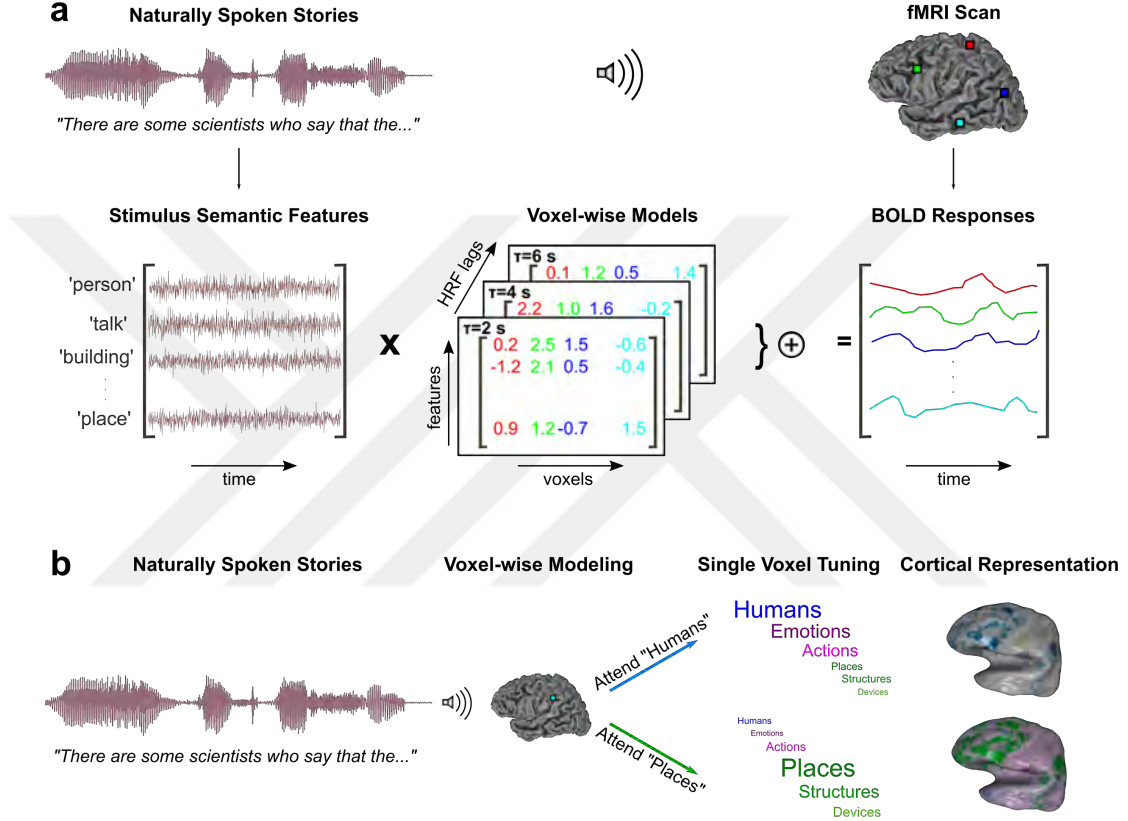


Figure 2.2: **Natural semantic search in spoken narratives.** (a) During the semantic search experiment, five subjects listened to three naturally spoken stories while attending to words that belonged to either one of two target categories: humans or places. Whole-brain BOLD responses were acquired. To estimate semantic tuning in single voxels, we fit voxelwise models that optimally predict measured BOLD responses in terms of the semantic features of the story stimuli. To account for hemodynamic delays, a finite impulse response filter was used for each semantic feature. The voxelwise tuning vectors were then averaged across different time delays. The resulting semantic tuning vectors describe how words in the story stimulus evoke BOLD responses. (b) Semantic tuning was estimated separately for each attention task. Tuning changes in single voxels are a unique, diagnostic aspect of the tuning-shift hypothesis. To test this hypothesis, we measured changes in voxel tuning between the two attention tasks.

## 2.2 Methods

### 2.2.1 MRI Protocols

MRI data were previously collected by researchers from The Gallant Lab in a 3T Siemens Tim Trio scanner at the University of California, Berkeley using a 32-channel head coil. Data from the passive listening experiment were used in a previous study [20] (see 2.2.3). Gradient EPI sequence was used with repetition time = 2.00 s, echo time = 33 ms, flip angle =  $70^\circ$ , voxel size =  $2.24 \times 2.24 \times 4.13$  mm<sup>3</sup>, slice thickness = 3.5 mm with 18% slice gap, field of view =  $224 \times 224$  mm<sup>2</sup> and 32 axial slices to cover the entire cortex. Anatomical data were collected using T1-weighted multi-echo MP-RAGE sequence with voxel size =  $1 \times 1 \times 1$  mm<sup>3</sup> and field of view =  $25 \times 212 \times 256$  mm<sup>3</sup>.

### 2.2.2 Subjects

Functional data were collected from five healthy adult volunteers (one female, four males). All subjects had normal hearing. The experimental protocol was approved by the Protection of Human Subjects Committee at the University of California, Berkeley and written informed consent was obtained from all subjects.

### 2.2.3 Experimental Procedure

In the passive listening experiment, subjects were instructed to listen to naturally spoken stories without engaging in any other task. The stimuli for passive listening experiment consisted of ten stories selected from The Moth Radio Hour which covers a wide range of topics: ‘Alternate Ithaca Tom’ by Tom Weiser; ‘Targeted’ by Jen Lee; ‘My Avatar and Me’ by Laura Albert; ‘My Unhurried Legacy’ by Kyp Malone; ‘Ode to Stepfather’ by Ethan Hawke; ‘Under the Influence’ by Jeffery Rudell; ‘How to Draw a Nekkid Man’ by Tricia Rose Burt; ‘My First Day at the

Yankees’ by Matt McGough; ‘Naked’ by Catherine Burns; and ‘Life Flight’ by Kimberly Reed. In each story, a speaker tells a memoir to a live audience. Passive listening data were collected in two separate 2-hour scanning sessions performed on different days.

In the attention experiment, subjects were instructed to listen to each story twice and search for words that are semantically related to humans (e.g., student, guy, parent, everyone) in one run and related to places (e.g., school, city, north, lake) in the other run. The stimuli for the attention experiment included three stories from The Moth Radio Hour (‘From Boyhood to Fatherhood’ by Jonathan Ames, ‘My First Day at the Yankees’ by Matt McGough, and ‘Naked’ Catherine Burns). Attention data were collected in a single scan session consisting of six runs. The target category was fixed in each run. The attention tasks were alternated in consecutive runs. To ensure continuous vigilance, subjects were cued to depress a button whenever they hear a target word. All stimuli were played at 44.1 kHz using Sensimetrics S14 in-ear piezoelectric headphones. Frequency response was flattened using hardware parametric equalizer and volume was normalized to have peak loudness of -1 dB relative to maximum.

## 2.2.4 ROI Abbreviations

Regions of interest used in our study are: Auditory cortex (AC), Broca’s area (Broca), Wernicke’s area (WER), angular gyrus (AngG), supramarginal gyrus (SupMG), frontal operculum (FO), intraparietal sulcus (IPS), frontal eye fields (FEF), superior/middle/inferior frontal gyri (SFG, MFG, IFG), retinotopic area (RET, consisting of V1, V2, V3, V4, LO, V3A, and V3B), fusiform face area (FFA), occipital face area (OFA), inferior frontal sulcus face patch (IFSFP), extrastriate body area (EBA), parahippocampal place area (PPA), retrosplenial cortex (RSC), occipital place area (OPA), supplementary eye fields (SEF), primary motor and somatosensory areas (S1F, M1F, S1H, M1H, S1M, and M1M), supplementary motor hand and foot areas (SMHA, SMFA).

### 2.2.5 Functional Localizers

We defined boundaries of common functional ROIs in each subject by using functional localizers, MNI template back transformation or anatomical inspection. Functional localizers include a visual category localizer, an auditory cortex localizer, a retinotopic localizer, and a motor localizer.

In the visual category localizer, the subject was shown either places, faces, human body parts, animals, household objects, or spatially scrambled household objects during six 4.5-minute scans. Fusiform face area (FFA), occipital face area (OFA), inferior frontal sulcus face patch (IFSFP) were defined based on the faces versus objects contrast [55]. Extrastriate body area (EBA) was defined based on the human body parts versus objects contrast [56]. Parahippocampal place area (PPA), retrosplenial cortex (RSC) and occipital place area (OPA) were defined based on the scenes versus objects contrast [57, 58] .

The auditory cortex localizer data were collected in one scan. Subjects listened to 10 repeats of a 1-minute auditory stimulus consisting of music, speech, and natural sounds. F-statistic of each voxel’s response across 10 repeats was used to define the auditory cortex (AC).

Retinotopic mapping data were collected in four 9-minute scans. Clockwise and counterclockwise rotating wedges, and expanding and contracting rings were shown to the subject. Using visual angle and eccentricity maps, V1, V2, V3, V4, LO, V3A, and V3B were defined. A broad retinotopic area (RET) was defined based on V1, V2, V3, V4, LO, V3A, and V3B in ROI based analyses.

Motor localizer data were collected in a single 10-minute scan. The subject performed five different motor tasks in a random order. The subject performed either small finger movements (‘hand’ cue), small toe movements (‘foot’ cue), small mouth movements (‘mouth’ cue), self-generated speech (‘speak’ cue), or saccade movements (‘saccade’ cue). Frontal eye fields (FEF), intraparietal sulcus (IPS), frontal operculum (FO) and supplementary eye fields (SEF) were defined based on the saccade versus rest contrast. Primary motor and somatosensory

areas (S1F, M1F, S1H, M1H, S1M, and M1M) and supplementary motor hand and foot areas (SMHA, SMFA) were defined based on the foot, hand, and mouth versus rest contrast.

Additional ROI definitions were obtained for language areas based on the MNI template. To do this, the MNI coordinates of Brodmann areas were back-transformed to the brain spaces of individual subjects. Angular gyrus (AngG, Brodmann area (BA) 39), supramarginal gyrus (SupMG, BA 40), Wernicke’s area (WER, between BA 39-40), anterior cingulate cortex (ACC, BA 32-33), posterior cingulate cortex (PCC, BA 23-31), superior/middle/inferior frontal gyri (SFG, MFG, IFG) were defined using MNI template. When needed, we fine-tuned these definitions by inspecting anatomical landmarks. ACC and PCC were fine-tuned using the location of cingulate sulcus. SupMG, WER and AngG were fine-tuned using the posterior endpoints of superior temporal sulcus and lateral sulcus. SFG, MFG and IFG were fine-tuned using the locations of superior frontal sulcus and inferior frontal sulcus.

## 2.2.6 Data Pre-processing

Each functional run was motion corrected using the Motion Correction FMRIB Linear Image Registration Tool (MCFLIRT) in FSL 5.0 [59]. Functional images within each run were first registered to the temporal mean image across that run. To align all runs to the same reference, the temporal means were registered to a pre-selected image of each subject. These two transformations were concatenated, and applied to functional images in a single step. The resulting motion-corrected time series data were manually checked for accuracy.

Following motion correction, low frequency drifts in BOLD responses were removed by detrending. The drifts were estimated using a 3rd order Savitsky-Golay filter with a 240-second window, and then removed from measured BOLD responses.

To extract cortical voxels, the Brain Extraction Tool in FSL was used. Cortical

voxels were defined as the set of voxels within an 8-mm radius of the cortical surface. There were total of  $75,286 \pm 6,887$  cortical voxels (mean  $\pm$  s.d. across subjects). Subsequent analyses were restricted to the cortical voxels identified in each subject. During ROI analyses, non-cortical voxels were further excluded, thus selection of a relatively large radius for cortical sheet does not constitute a problem.

### 2.2.7 Head Motion and Physiological Regressors

In order to minimize bias from head motion or physiological noise, we used estimates of these factors as nuisance regressors during model fitting. Head motion regressors were simply the translational and rotational variations in head position calculated during motion correction yielding a total of six regressors. During the experiments, the cardiac and respiratory states were recorded using a pulse oximeter and a pneumatic belt. Raw outputs of the pulse oximeter and pneumatic belt were the cardiac state diagram and thoracic modulation diagram, respectively. Due to low SNR, only the first few FS components explain meaningful variations in physiological noise. Thus, using the PhLEM Toolbox, we computed the first two FS components of the cardiac state and the first FS component of the respiratory state [60]. Furthermore, we computed the instantaneous heart rate per TR from cardiac state yielding a total of seven physiological regressors. Both motion and physiological regressors were included as nuisance regressors during model fitting. The inclusion of these nuisance regressors did not affect the model performance in any brain region where the semantic model provided significant response predictions.

### 2.2.8 Semantic Model Construction

To estimate semantic tuning of single voxels, we constructed a semantic model that explains BOLD responses in terms of the semantic features of the story stimuli. To extract semantic features, a word embedding space was formed based

on word co-occurrence statistics. The assumption here is that words with similar meaning tend to occur in nearby positions in text. To estimate word co-occurrences, a large corpus of text was compiled. This corpus contained transcripts of all Moth stories (including the stories used in the experiment), popular books, Wikipedia pages, and user comments gathered from reddit.com. We first constructed a lexicon of 10,470 words by combining the set of all words in the stories with the 10,000 most common words in the corpus. We then calculated co-occurrences between 985 basis words (from Wikipedia’s List of 1000 Basic Words consisting of common words from a very broad range of topics) and the 10,470 lexicon words. A single co-occurrence was taken as the appearance of a lexicon word within 15-word neighborhood of the unique word. This procedure yielded a co-occurrence matrix of size  $985 \times 10,470$ . The co-occurrence matrix was first log-transformed, then its rows were z-scored to account for differences in the frequency of unique words, and its columns were z-scored to account for differences in the frequency of lexicon words. Each column of the resulting co-occurrence matrix corresponds to the projection of a lexicon word onto the embedding space. Thus, each lexicon word is represented by a 985-dimensional semantic feature vector. Additionally, to rule out possible biases due to correlations between semantic features and word-rate, phoneme rate, and phoneme types, we included these features as nuisance regressors in our semantic model. To project the story stimuli onto the word embedding space, the procedures in Huth et. al. were followed [20]. The time course of the feature vectors were resampled using a 3-lobe Lanczos filter with a cut-off frequency of 0.249 Hz (the Nyquist frequency of the fMRI acquisition). This procedure yielded a stimulus matrix of  $985 \times 1128$  (Number of features  $\times$  Number of TRs).

## 2.2.9 Voxelwise Model Estimation and Validation

Voxelwise models were fit between the stimulus matrix and BOLD responses using  $L_2$ -regularized ridge regression. To account for hemodynamic delays, a finite impulse response filter with four time delays (1, 2 and 3 TRs) was applied to each model feature. For BOLD responses in each time instant (TR), we used

feature set of the story stimulus at 1 TR, 2 TR, and 3 TR before the time instant. Thus, for each TR, we have stimulus feature vector of size  $3 \times 985$  ( $3 \times \text{Number of Features}$ ). Final voxelwise tuning vectors were taken as the dot product of the resulting model weights with the weights of the linear temporal filter.

Model fitting and performance evaluations were implemented via a nested cross-validation procedure. To measure prediction performance, an outer 500-fold cross-validation was used. In each outer fold, the total of 1128 time samples were split into a block of 900 samples for model fitting and a remaining block of 228 test samples. To fit models, an inner 12-fold cross-validation was performed within the samples reserved for model fitting (900 samples). In each inner fold, this data were further split into 700 training samples and 200 validation samples for selection of regularization parameters ( $\lambda$ ). Separate models were fit for 30 different regularization parameters in the range 10 to 1000 (spaced logarithmically). Prediction scores were taken as the Pearson’s correlation between measured and predicted BOLD responses. The regularization parameter that maximized the average prediction score across inner folds was selected for each voxel separately. Based on the optimal regularization parameter, the prediction score of each voxel was taken as the average prediction score measured across outer folds.

To keep the scale of the weights consistent in subsequent analyses, a single value of the regularization parameter was used for all voxels in all subjects. This regularization parameter was selected by maximizing the average prediction scores of cortical voxels pooled across subjects ( $\lambda^* = 204.34$ ). Final voxelwise models based on  $\lambda^*$  were refit using all 1128 time samples in the experimental dataset.

### **2.2.10 Construction of the Semantic Space**

To obtain a common semantic space across subjects, we implemented principal component analysis (PCA) on the tuning vectors measured during the passive listening task. To avoid noise from poorly modeled voxels, we selected the 10,000 best modeled voxels from each subject and obtained 985 PCs. To maximize the quality of the semantic space, we selected only the first three PCs that captured

$32.25 \pm 2.33\%$  (mean  $\pm$  s.d. across subjects) of the variance in tuning vectors. We assigned RGB colors to the voxels. Red, blue and green channels of each voxel were assigned according to the loadings of voxelwise tuning vector on the first, second and third PCs. We then visualized semantic tuning of cortical voxels on flat maps under each attention task separately. To help interpretation of these maps, we also projected the words in the stories onto the three PCs. We identified six semantic word categories via k-means clustering. We manually labeled each category. Each word had a 985-dimensional semantic feature vector in the embedding space. We projected the feature vector of the words onto the three PCs. We then assigned RGB color to each word according to the projections of its feature vector on the first, second and third PCs. Finally, the color of each semantic word category was taken as the average RGB color of the words within that category.

### **2.2.11 Tuning Shift Index**

Attentional tuning shifts towards to target word category will increase the semantic selectivity for the target. Thus, tuning shifts of all cortical voxels can be assessed by quantifying the tuning strengths for humans and places separately during different attention tasks. Tuning strength of a voxel for humans was defined as the similarity between that voxel’s semantic tuning vector and an ideal human template. Similarly, tuning strength of a voxel for places was defined as the similarity between that voxel’s semantic tuning vector and an ideal place template. The templates were taken as tuning vectors of a voxel tuned solely for humans or places. To construct the templates, the feature vectors of the set of basic words that belong to the target categories were averaged in the embedding space. For the human template, this set included words that are closely related to humans (i.e., human, person, people, someone, and man). For the place template, this set included words that are closely related to places (i.e., place, building, location, area, and city). The tuning strength for a target category was then quantified as the Pearson’s correlation between the voxelwise tuning vector

and the respective template:

$$\begin{aligned} s_{i,H} &= \text{corr}(w_i, t_H) \\ s_{i,P} &= \text{corr}(w_i, t_P) \end{aligned} \tag{2.1}$$

Here,  $w_i$  is the voxel tuning vector while attending to category- $i$ ,  $i \in (H,P)$ ,  $t_H$  and  $t_P$  are the ideal human and place templates.  $s_{H,H}$  is the tuning strength for humans when attending to humans and  $s_{P,H}$  is the tuning strength for humans when attending to places. Meanwhile,  $s_{P,P}$  is the tuning strength for places when attending to places and  $s_{H,P}$  is the tuning strength for places when attending to humans. The tuning shift index (TSI) was finally calculated as [53]:

$$TSI = \frac{(s_{H,H} - s_{H,P}) + (s_{P,P} - s_{P,H})}{2 - \text{sign}(s_{H,H} - s_{H,P}) s_{H,P} - \text{sign}(s_{P,P} - s_{P,H}) s_{P,H}} \tag{2.2}$$

The numerator sums the differences in tuning strengths for the target versus non-target categories across attention tasks. The denominator scales the TSIs to range between -1 and 1. Tuning shifts in the direction of the attended category will yield positive TSIs while tuning shifts in the opposite direction of attended category will yield negative TSIs. A voxel that is ideally tuned to the attended category will have a TSI of 1 and a voxel that is tuned to the distractor category will have a TSI of -1. A voxel that did not shift its tuning between two attention tasks, will get a TSI of zero.

### 2.2.12 Control Analysis in the Absence of Target Stimuli

To ensure that tuning shifts are in the direction of attended categories even when no target is present, we performed a control analysis. We fit semantic models using only those segments of the story stimuli that did not contain any words semantically related to humans or places. We then performed TSI analyses on the fit models similarly to the main analysis. Model estimation and subsequent analyses were then repeated on this reduced dataset.

### 2.2.13 Behavioral Analysis

In the attention experiment, subjects were cued to depress a button on a fiber-optic response pad whenever they heard a word that belong to the target category in the story. To calculate the success rate, a ‘hit’ was defined as a button press detected within one TR of target appearance in the stories, and a ‘false alarm’ was defined as a button press detected in the absence of a target. We found that the subjects performed consistently well with  $87.55 \pm 8.23\%$  success rate and  $14.91 \pm 6.05\%$  false alarm rate (mean  $\pm$  s.d.). The behavioral performance, measured by the sensitivity index ( $d'$ ), was compared across the two attention tasks using Wilcoxon rank-sum tests. Participants performed equally well when searching for either category, indicating that the task difficulty was balanced across attention tasks (Wilcoxon rank-sum test,  $P > 0.05$ ).

### 2.2.14 Significance of Attentional Modulations

If attentional tuning modulations are significant during story listening, the BOLD responses of a given task should be better predicted by the semantic model trained for that task compared to the semantic model trained for the other attention task. We first fit a self-prediction model for each voxel by training and testing the semantic model on the same attention task. We then fit a cross-prediction model by training the semantic model on one attention task and testing it on the other attention task. The prediction scores of the self- and cross-prediction models were compared using t-tests for each voxel (corrected for multiple comparisons via FDR). To measure the percentage of response variance that can be attributed to attentional modulations, we calculated the difference in variance explained by the self- and cross-prediction models.

### 2.2.15 Construction of Control Models

To ensure that semantic tuning cannot be attributed to selectivity for low-level features, we constructed two control models: a phoneme model and a spectrogram model [20]. The phoneme model comprised features including the number of phonemes that appear within each TR (1 feature), the number of words that appear within each TR (1 feature), and the number of occurrence for each phoneme within each TR (39 features). The spectrogram model features reflected time varying spectral power density of the speaker voice. We used the Welch method to estimate the time-varying power spectral density and averaged the power spectral density for 2 second segments (period of fMRI acquisition) [61]. To prevent large negative values, we limited the power spectrum to a 50 dB ceiling threshold. Finally, the spectrogram model consisted of a 300 dimensional vector that reflected the power of the signal between 0 Hz and 15 kHz in 50 Hz bands.

### 2.2.16 Cortical Flat Maps

Cortical flat maps for each subject were generated based on  $T_1$ -weighted anatomical scans using Freesurfer [62]. Prior to surface reconstruction, anatomical surface segmentations were manually corrected using pycortex [63]. Relaxation cuts were made into the surface of each hemisphere and the surface crossing the corpus callosum was removed. The calcarine sulcus cut was made at the horizontal meridian in V1. Using boundary based registration (BBR) in FSL, functional images were aligned to the cortical flat maps. The mid-cortical surface that lies halfway between the pial and white matter surfaces was selected for visualization.

Semantic model prediction performance maps in Supp. Figs. A.1a, A.2a, A.3a, A.4a, A.5a were created by projecting voxelwise prediction scores onto the cortical flat maps using nearest-neighbor interpolation. TSI flat maps in Supp Figs. A.1b, A.2b, A.3b, A.4b, A.5b were created based on the same procedure. To visualize TSI values, a color map was created where voxels with TSI less than -0.2 were assigned blue color while voxels with TSI greater than 0.2 were assigned red color.

Voxels with TSI of 0 were assigned gray color. Voxels with TSI between -0.2 and 0 were assigned linearly changing color between blue and gray. Voxels with TSI between 0 and 0.2 were assigned linearly changing color between gray and red.

## 2.3 Results

To estimate semantic tuning of single voxels, we used a voxelwise semantic model that explains BOLD responses in terms of the semantic features of the story stimuli. To extract semantic features, a word embedding space was formed based on word co-occurrence statistics in a large corpus of English text (see 2.2.8). The dimensions of the embedding space were selected as words from Wikipedia’s List of 1000 Basic Words (note that this list contains only 985 unique words). Each word was projected onto the embedding space by calculating the co-occurrence of that word with the 985 dimensions. Locations of the words in the embedding space reflect semantic relationships. Semantically similar words should project to nearby points while semantically distinct words should project to distant points. For example, ‘school’ and ‘class’ are nearby in semantic space (correlation distance,  $1 - r = 0.26$ ), whereas ‘school’ and ‘arm’ are far apart (correlation distance,  $1 - r = 1.22$ ). Therefore, semantic features of a given word were taken as its location in the embedding space. To find the semantic features present in the story stimuli, each word in the story transcripts was projected onto the word embedding space. Separately for each attention task, we then used ridge regression to estimate model weights that optimally predict measured BOLD responses in single voxels. The fit model weights reflect voxelwise semantic tuning as a 985-dimensional vector (Representative semantic tuning vectors of a single voxel are visualized in Fig. 2.3).

In a recent study, authors found that during natural story listening, the semantic model used here explains much of variance in BOLD responses of cortical voxels [20]. It is likely that the semantic model also explains a large amount of response variance during natural semantic search, as well. To test model performance, the fit models were used to predict BOLD responses in a held-out data

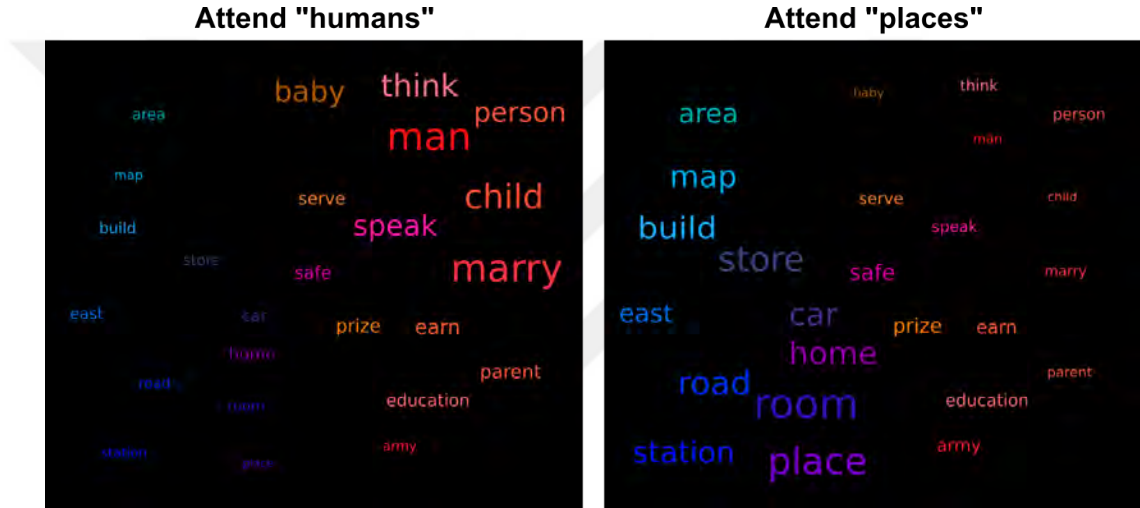


Figure 2.3: **Attentional tuning changes for a representative voxel in intraparietal sulcus.** Semantic tuning vectors of a well-modeled voxel within IPS, in subject S1, for the two attention tasks. To simplify visualization, only a subset of words that are either neutral, semantically related to humans, or semantically related to places are shown. The color and location of the words reflect their coordinates in the semantic space, e.g., red-magenta words are related to the humans category whereas blue-purple words are related to the places category (see 2.2.10). The size of each word represents the magnitude of the evoked BOLD response. This voxel has a prediction score of 0.401 and exhibits significant attentional modulation ( $t$  test,  $q(\text{FDR}) < 0.05$ ). The voxel is strongly tuned for words related to the target category in both tasks, and weaker tuning is observed for behaviorally irrelevant words.

set. Prediction score was taken as Pearson’s correlation between predicted and measured responses. For each voxel, a single prediction score was computed by averaging the scores across two attention tasks. Overall  $63.24 \pm 4.90\%$  (mean  $\pm$  s.d. across subjects) of cortical voxels are significantly predicted by the semantic model ( $t$  test,  $q(\text{FDR}) < 0.05$ ). Cortical distribution of prediction scores in individual subjects are shown in Supp. Figs. A.1a, A.2a, A.3a, A.4a, A.5a. We find that the semantic model accurately predicts cortical voxels in many areas including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC) and inferior prefrontal cortex (IPFC). These results suggest that semantic tuning vectors accurately reflect responses of many cortical voxels across temporal, parietal, and prefrontal cortices during natural semantic search.

In a recent study, the authors reported that attention to object categories during natural vision causes widespread modulations in category tuning [53]. Therefore, it is likely that attention to word categories during story listening also causes similar modulations in semantic tuning across cortex. To assess the significance of attentional modulations during natural semantic search, we performed an analysis based on prediction scores. We reasoned that if attentional tuning modulations are significant, the BOLD responses of a given attention task should be better predicted by the semantic model fit for that task compared to the semantic model fit for the other task (see 2.2.14). We first fit a self-prediction model for each voxel by training and testing the semantic model on the same attention task. We then fit a cross-prediction model by training the semantic model on one attention task and testing it on the other attention task. We compared the prediction scores of the self- and cross-prediction models. We find that  $57.80 \pm 4.90\%$  (mean  $\pm$  s.d. across subjects) of cortical voxels have significant tuning changes ( $t$  test,  $q(\text{FDR}) < 0.05$ ). Furthermore,  $24.35 \pm 1.60\%$  (mean  $\pm$  s.d.) of these voxels are well-modeled (prediction score was greater than 1 s.d. above the mean). These results indicate that many cortical voxels exhibit significant attentional modulation during natural semantic search.

In naturally spoken stories, there may be intrinsic correlations between high-level semantic features and low-level phonemic or spectrogram features. If such correlations are significant then they could bias the semantic models that are estimated here. Thus, it is possible that our results partly reflect tuning for low-level features in the story stimuli. To ensure that voxelwise semantic tuning cannot be attributed to selectivity for low-level features, we constructed two separate control models: a phoneme model that measures tuning for number of phonemes, number of words, and phoneme types, and a spectrogram model that measures tuning for time varying spectral power density of the speaker voice (see 2.2.8). Those models were separately fit to single voxels and prediction scores were calculated. The semantic model has significantly higher prediction scores than both the phoneme model and the spectrogram model in voxels that are well-modeled by the semantic model (Wilcoxon signed-rank test,  $P < 0.05$ . See Supp. Fig. A.6 for prediction scores of the semantic, phoneme and spectrogram models in classical functional regions of interest (ROIs)). The percentage of response variance explained in well-modeled voxels is  $15.35 \pm 3.88\%$  (mean  $\pm$  s.d. across subjects) for the semantic model,  $4.80 \pm 0.91\%$  (mean  $\pm$  s.d. across subjects) for the phoneme model, and  $3.54 \pm 0.46\%$  (mean  $\pm$  s.d. across subjects) for the spectrogram model. Furthermore, the percentage of response variance in these voxels that can be attributed to attentional modulations is  $7.06 \pm 3.10\%$  (mean  $\pm$  s.d. across subjects) for semantic tuning and only  $0.68 \pm 0.84\%$  for phonemic tuning,  $-0.03 \pm 0.22\%$  for spectro-temporal tuning. Thus, the attentional modulations that we observe in semantic tuning cannot be attributed to tuning for low-level features including phoneme or spectrogram features of the story stimuli.

### 2.3.1 Cortical Distribution of Attentional Modulations

To assess the nature and distribution of attentional modulations, we separately visualized voxelwise semantic tuning for each task on cortical flat maps. We projected semantic tuning vectors onto a common semantic space across subjects. The semantic space was obtained by principal component analysis (PCA) on tuning vectors measured during a separate passive listening task (see 2.2.10).

The first three PCs were selected that capture  $32.25 \pm 2.33\%$  (mean  $\pm$  s.d. across subjects) of variance in tuning vectors. To help interpretation, we also projected all the words in the story stimuli onto the same PCs and identified six semantic word categories by k-means clustering (see 2.2.10). Cortical flat maps of semantic tuning for passive listening and the two attention tasks are shown in Fig. 2.4 for a representative subject (see Supp. Figs. A.7, A.8, A.9, A.10, A.11 for individual subjects). During search for humans, many voxels in posterior areas that are tuned for places under passive listening (e.g., cingulate sulcus, IPS) shift their tuning away from places. Voxels that are not tuned for humans under passive listening (in some parts of frontal gyrus, and FEF) shift their tuning toward humans. During search for places, many voxels in posterior areas that are tuned for humans under passive listening (for example, cingulate sulcus) shift their tuning away from humans, and voxels that are not tuned for places under passive listening (in frontal gyrus, and some parts of precentral sulcus) shift their tuning toward places. No visible changes in semantic tuning are observed in voxels across auditory and visual cortices. These results imply that semantic attention causes modulations in semantic tuning within many areas that are implicated in language perception and attentional control during natural semantic search.

### **2.3.2 Attention Shifts Semantic Tuning towards the Target Category**

Next, we asked whether the attentional modulations in semantic tuning that we measured are consistent with the tuning-shift hypothesis. The first prediction of the tuning-shift hypothesis is that attention expands the semantic representation of target word categories by causing tuning shifts in cortical voxels toward the target. To test this prediction, we visualized the distribution of tuning across well-modeled voxels. For each task, semantic tuning vectors were projected onto the semantic space. In this space, voxels with similar semantic tuning will project to nearby points whereas voxels with different semantic tuning will project to distant points (Fig. 2.5a). We find that most well-modeled voxels are selectively tuned

for the attended category, and attention causes tuning shifts in many voxels.

To quantify the degree of shifts towards or away from the target word categories, we next performed a tuning shift analysis. For this purpose, we measured the tuning of individual voxels to target categories. Idealized human and place templates were created to represent voxels that are tuned solely for the target categories (see 2.2.11). For each task, voxelwise selectivity for a target category was taken as the Pearson’s correlation between the semantic tuning vector and the corresponding idealized template. To estimate the magnitude and direction of tuning shifts, we computed a tuning shift index (TSI) that summarizes the difference in selectivity for the attended versus unattended category (see 2.2.11). A voxel that shifts its tuning toward attended category will have a positive TSI, while a voxel that shifts its tuning away from attended category will have a negative TSI. We find that the average TSI of well-modeled voxels is significantly greater than 0 in all subjects (Wilcoxon signed-rank test,  $P < 0.05$ ; Fig. 2.5c and Supp. Fig. A.12). Prior to TSI calculations, all responses were z scored individually within each attention task, so tuning shifts cannot be attributed to changes in voxel response baseline or gain (see 2.3.5). Hence, these results are in agreement with the prediction that semantic search expands the semantic representation of the attended category across cortex.

Previous fMRI and neurophysiology studies on vision report that visual attention causes widespread tuning shifts across cortex [41, 64, 65]. It is therefore likely that tuning shifts during natural semantic search are also widespread. To visualize the cortical distribution of tuning shifts, we projected TSI values onto cortical flat maps (Fig. 2.5b and Supp. Figs. A.1b, A.2b, A.3b, A.4b, A.5b). Voxel clusters with positive TSI are observed in many ROIs implicated in attention control including frontal operculum (FO), frontal eye fields (FEF), intraparietal sulcus (IPS), in language perception including Broca’s area (Broca), Wernicke’s area (WER), supramarginal gyrus (SupMG), and other areas including precen-tral sulcus and cingulate sulcus. In contrast, voxel clusters with negative TSI are observed in posterior cingulate cortex (PCC), angular gyrus (AngG) and right superior frontal gyrus (SFG) that partly overlap with the default mode network [66]. This result suggests that the default mode network areas play a role in

distractor suppression and error monitoring during semantic search tasks.

To examine how specific cortical areas alter their selectivity for attended and unattended categories, we performed detailed analyses of tuning shifts in several common ROIs (Fig. 2.6, see Supp. Fig. A.13 for individual subjects). Consistently across subjects, many areas in inferior parietal lobe and frontal gyrus including Wernicke’s area (WER), supramarginal gyrus (SupMG), frontal operculum (FO), intraparietal sulcus (IPS), angular gyrus (AngG), frontal eye fields (FEF), superior frontal gyrus (SFG), anterior cingulate cortex (ACC), middle and inferior frontal gyri (MFG, IFG) have significantly positive TSI (Wilcoxon signed-rank test,  $P < 0.05$ ). TSI is relatively lower in auditory cortex (AC), but is increased towards the later stages of language processing (WER and AngG). Largest TSI values are observed in frontal areas that take part in language comprehension (Broca’s area) and attention control (FO, FEF and IPS). The relatively modest level of tuning shift in RET can be attributed to imagery [67]. These results imply that attentional tuning shifts become progressively stronger toward later stages of language processing and attention control.

Left hemisphere dominance has been reported during language comprehension tasks in areas along the frontal gyrus and superior temporal gyrus [68, 69]. Recent studies on auditory attention further report that response modulations exhibit left hemisphere dominance in AC [70]. It is thus likely that tuning shifts during natural semantic search are also dominant in the left hemisphere. To test for hemispheric dominance, we compared the TSI values in each ROI across hemispheres. We find that TSI values in many areas implicated in language comprehension (SupMG, SFG, MFG, PCC) and attention control (FO, FEF, IPS) are significantly higher in the left hemisphere (Mann-Whitney U test,  $P < 0.05$ ). TSI values are also higher in the left hemisphere for Broca and IFG for all subjects except S3. These results indicate that attentional modulations in semantic tuning are left lateralized broadly across cortex.

### 2.3.3 Attention Shifts Semantic Tuning in the Absence of Targets

The second prediction of the tuning-shift hypothesis is that attention changes semantic tuning of cortical voxels even when no targets are present in the stimuli. To test this prediction, we fit semantic models using only those segments of the story stimuli that did not contain any words semantically related to humans or places. We then performed TSI analyses on the fit models similarly to the main analysis. Note that this analysis serves as an additional control against nuisance variations across attention tasks including differences in heart rate and respiration due to target detection. We find that the average TSI across well-modeled voxels is significantly greater than zero in each individual subject (Wilcoxon signed-rank test,  $P < 0.05$ , Supp. Fig. A.12 b). Moreover,  $73.74 \pm 2.19\%$  (mean  $\pm$  s.d. across subjects) of cortical voxels have the tuning shifts in consistent directions across the main and control analyses. These results indicate that high-level auditory attention shifts semantic tuning of most cortical voxels towards the attended category even in the absence of targets.

### 2.3.4 Attention Warps Semantic Representation

A recent fMRI study on natural story listening suggest that the human brain represents words by embedding them into a low dimensional semantic space whose dimensions are represented continuously across cortex [20]. The third prediction of the tuning shift hypothesis is that during semantic search attention should warp the semantic space to expand the representation of target categories while compressing the representation of behaviorally irrelevant words. To test this prediction, we separately estimated the topology of the semantic space under passive listening and the two attention tasks. For this purpose, PCA was performed on semantic tuning vectors estimated independently for each task (see 2.2.10). The subset of 100 most common words in English were selected to visualize the semantic spaces (Fig. 2.7). During passive listening, the selected words show a relatively uniform distribution across the semantic space. During attention to

humans, representations of words semantically similar to humans (e.g., person, human, family, and children) are expanded while representations of words semantically similar to places (e.g., home, place, road, and building) are compressed. During attention to places, representations of words semantically similar to places are expanded while representations of words semantically similar to humans are compressed. To quantify the degree and direction of warping in semantic space, we computed a semantic warping index (see 2.2.10). If the target words are more broadly spread in the semantic space compared to non-target words, the semantic warping index will have a value greater than 1. If the target words are less broadly spread in the semantic space compared to non-target words, the semantic warping index will have a value smaller than 1. The semantic warping index is significantly greater than 1 across subjects (Wilcoxon signed-rank test,  $P < 0.05$ ). These findings indicate that tuning shifts in cortex warp semantic representation in favor of attended words while compressing the representation of irrelevant words.

### 2.3.5 Discussion

Our results indicate that semantic search during natural story listening causes semantic tuning changes that cannot be explained by response baseline or gain changes in single voxels. These tuning changes warp the representation of both attended and unattended word categories across cortex. Moreover, tuning shifts occur even when words belonging to target categories are not present in the story. These results are consistent with an attentional mechanism that expands the cortical representation of word categories relevant to the target and compresses the representation of behaviorally irrelevant categories.

Several neuroimaging studies have examined the effects of high-level auditory attention on population responses in the human brain [38, 39, 40]. Response baseline changes were observed in temporal gyrus, frontal gyrus, and prefrontal cortex. However, these studies did not explore voxel-based tuning and could not distinguish voxel-based changes in tuning from changes in response baseline or

gain. Here, we instead used complex natural story stimuli that enabled us to measure responses to thousands of different words. We normalized BOLD responses of individual voxels in each task prior to model fitting to remove attentional effects on response baseline and gain. Lastly, we used a powerful voxelwise modeling framework to measure semantic tuning profiles across 1463 distinct words in our natural story stimuli. For the first time in literature, this experimental approach has enabled us to sensitively detect semantic tuning changes in single voxels.

Previous studies on low-level auditory attention tasks reported significant tuning changes in STRFs within primary auditory cortex [45, 47, 46, 48, 49, 50, 51, 52]. Thus one might question whether semantic tuning changes reported here are biased by shifts in tuning for low-level features (e.g., phonemic or spectro-temporal features). To rule out this bias, we performed control analyses using two separate models: the phoneme model to measure tuning for phonemic features and the spectrogram model to measure tuning for spectro-temporal features. When compared against these control models, the semantic model explains substantially more variance in BOLD responses across all examined ROIs. Furthermore, the control models do not explain the attentional response modulations in voxels that are significantly predicted by the semantic model. Therefore changes in semantic tuning measured here cannot be attributed to tuning changes for low-level features captured by the phonemic and spectro-temporal models. One can also wonder whether the attention task in the experiment is purely semantic or not. In fact, the task contains joint motor and mental tasks. For example, the subjects are supposed to press a button whenever they detect a target word. To prevent the potential biases due to the motor task or due to any physiological effect on subjects succeeding target appearance, we designed a separate control analysis (2.2.12) and we verified that the results did not change significantly when compared to the main analysis (2.3.3). These results together suggest that the observed results are mostly due to changes in semantic tuning of cortical voxels.

Another important question to be answered is the role of bottom-up processing versus top-down feedback in measured tuning shifts. Here, we used the same story stimuli during the two separate tasks, thus any attentional tuning changes between two tasks should be a result of top-down feedback mechanisms in the

brain. We observed small tuning shifts in early auditory areas and more substantial tuning shifts towards higher brain areas within temporal and prefrontal cortices. Control analysis indicates that tuning shifts in these higher areas cannot be attributed to tuning for low-level stimulus features represented in early areas. This implies that attentional modulations primarily warp semantic representation at later stages of auditory processing. However, the slow nature of BOLD responses makes it difficult for any fMRI study to measure the temporal relationship between signals arising in different brain areas at these later stages of processing.

Because we measured BOLD responses, we cannot make direct inferences about the underlying neural mechanisms that mediate tuning shifts in single voxels. There may be several possible neural mechanisms that might contribute to voxel-wise tuning changes. One possibility is that response baseline or gain changes in neurons that are tuned to the target word categories contribute to tuning changes in single voxels. Thus, future neurophysiology studies are needed to dissociate the precise neural mechanisms and their relative contributions to the tuning shifts reported here.

Instead of voxelwise modeling, we may have used Multivariate pattern analysis (MVPA). MVPA is also being commonly used in modeling functional characteristics of cortical areas [71, 72]. However, unlike voxelwise modeling, in MVPA, BOLD responses from a cluster of voxels are jointly analyzed to assess functional characteristics of these clusters depending on task conditions. Thus, resulting models may offer suboptimal sensitivity in reflecting individual functional characteristics of single voxels. Specifically in this study, for constructing the semantic space that reflects the representation of language across the human brain, fine-grained tuning information is needed. Thus, we decided to use voxelwise modeling.

In conclusion, we find that natural semantic search for a word category warps the auditory semantic representation across cortex, expanding the representation of behaviorally relevant semantic categories at the cost of irrelevant ones. This result supports the idea that attention invokes a matched-filter that dynamically

warps representations depending on behavioral goals. This mechanism likely improves task performance during natural auditory semantic search by more effectively dedicating limited neural resources to the target word categories. Overall, these findings help explain the human ability to perform complex auditory tasks in an ever-changing natural environment.



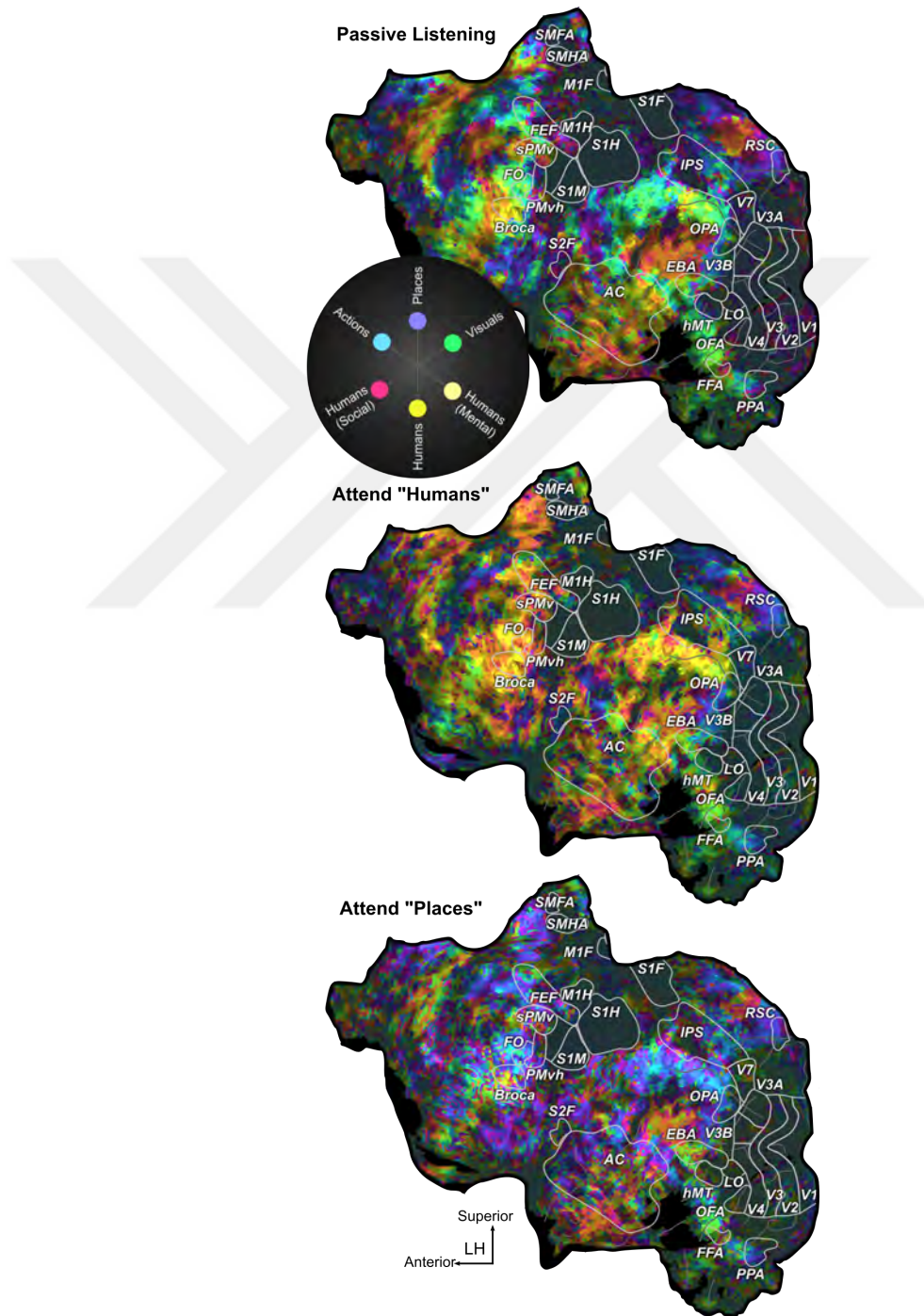


Figure 2.4: Caption next page.

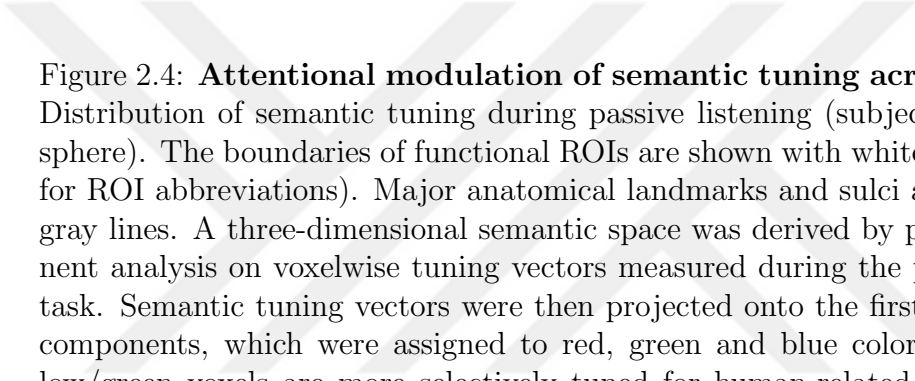


Figure 2.4: **Attentional modulation of semantic tuning across cortex.** (a) Distribution of semantic tuning during passive listening (subject S1, left hemisphere). The boundaries of functional ROIs are shown with white lines (See 2.2.4 for ROI abbreviations). Major anatomical landmarks and sulci are marked with gray lines. A three-dimensional semantic space was derived by principal component analysis on voxelwise tuning vectors measured during the passive listening task. Semantic tuning vectors were then projected onto the first three principal components, which were assigned to red, green and blue color channels. Yellow/green voxels are more selectively tuned for human-related words, whereas blue/purple voxels are more selectively tuned for place-related words. Regions of signal dropout, motor areas and insignificant voxels are shown in gray. (b) Distribution of semantic tuning during search for humans. While attending to humans, many voxels in posterior areas that are tuned to places under passive listening (e.g., Cingulate Sulcus, IPS) shift their tuning away from places. Voxels that are not tuned for humans under passive listening (e.g., in Precentral Sulcus, Frontal Gyrus, and FEF) shift their tuning toward humans. (c) Distribution of semantic tuning during search for places. While attending to places, many voxels in posterior areas that are tuned to humans under passive listening (e.g., Cingulate Sulcus) shift their tuning away from humans, and voxels that are not tuned for places under passive listening (e.g., in Frontal Gyrus, and Precentral Sulcus) shift their tuning toward places. These results indicate that many cortical voxels shift their semantic tuning toward target word categories.

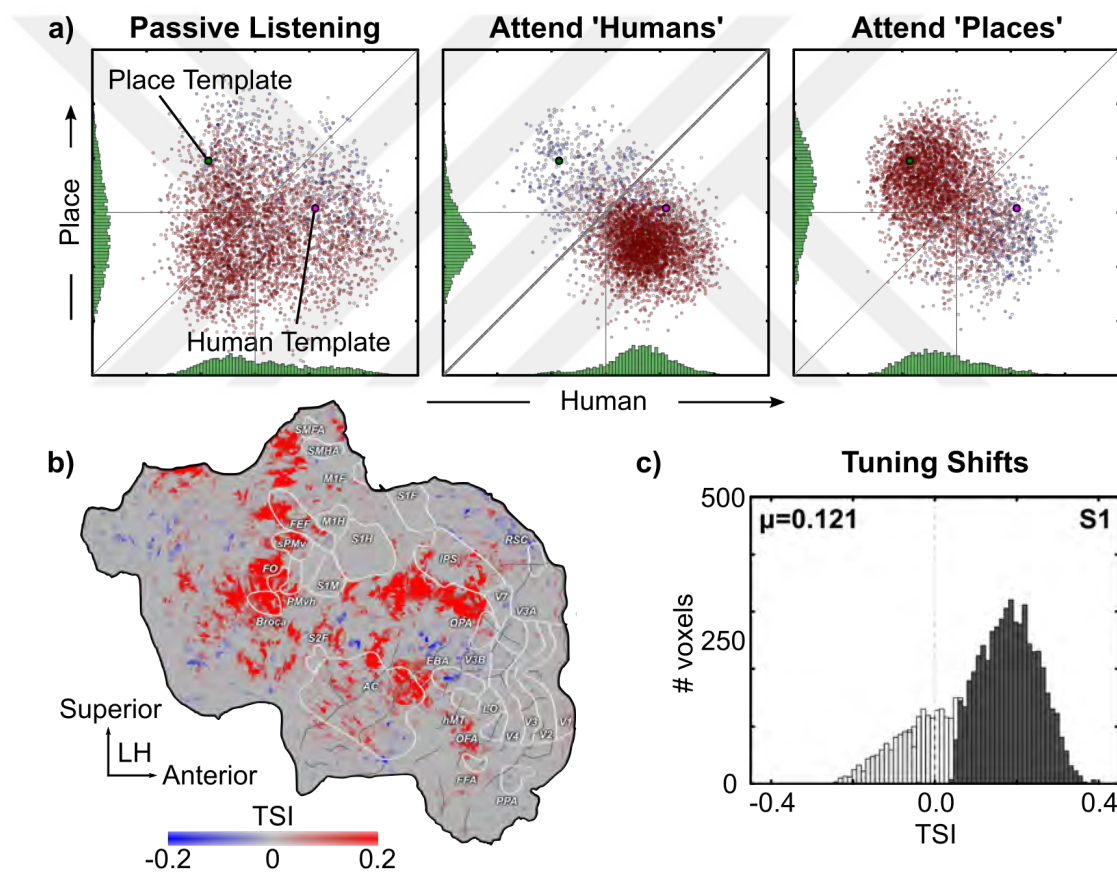


Figure 2.5: Caption next page.

Figure 2.5: **Attention shifts tuning of single voxels.** (a) To assess attentional changes in tuning of single voxels, we projected tuning vectors onto the first two dimensions of a semantic space. The semantic space was derived via principal component analysis (PCA) of voxelwise tuning vectors measured in a separate passive listening task. Only well-modeled voxels are shown (which exhibit significant attentional modulation and have prediction score greater than 1 s.d. above the mean). Each voxel is represented with a dot whose color indicates the tuning-shift index (TSI; see color bar in (b)). The projections of idealized human and place templates onto the semantic space are shown with two large dots. The marginal distributions of selectivity for humans and places are displayed with histograms in green. Many well-modeled voxels shift their tuning toward the target categories (Wilcoxon signed-rank test,  $q(\text{FDR}) < 0.05$ ). (b) Voxelwise TSIs were projected onto the cortical flat map of the left hemisphere for subject S1. Voxels in many brain areas implicated in attention control including frontal operculum (FO), frontal eye fields (FEF), and intraparietal sulcus (IPS), in language comprehension including Broca’s area (Broca), Wernicke’s area (WER), supramarginal gyrus (SupMG), and inferior frontal gyrus (IFG) shift their tuning toward the target category. In contrast, voxels in posterior cingulate cortex (PCC) and angular gyrus (AngG) which partly overlap with the default mode network shift their tuning away from the target. This result suggests that the default mode network areas play a role in distractor suppression and error monitoring during semantic search tasks. (c) Histogram of the TSIs for well-modeled voxels (8401 voxels). Significantly positive TSIs are represented by dark gray bars and significantly negative TSIs are represented by light gray bars (Wilcoxon signed-rank test,  $q(\text{FDR}) < 0.05$ ). The average TSI value is significantly greater than 0 (Wilcoxon signed-rank test,  $P < 0.05$ ). These results indicate that attention shifts tuning of most well-modeled voxels toward the target.

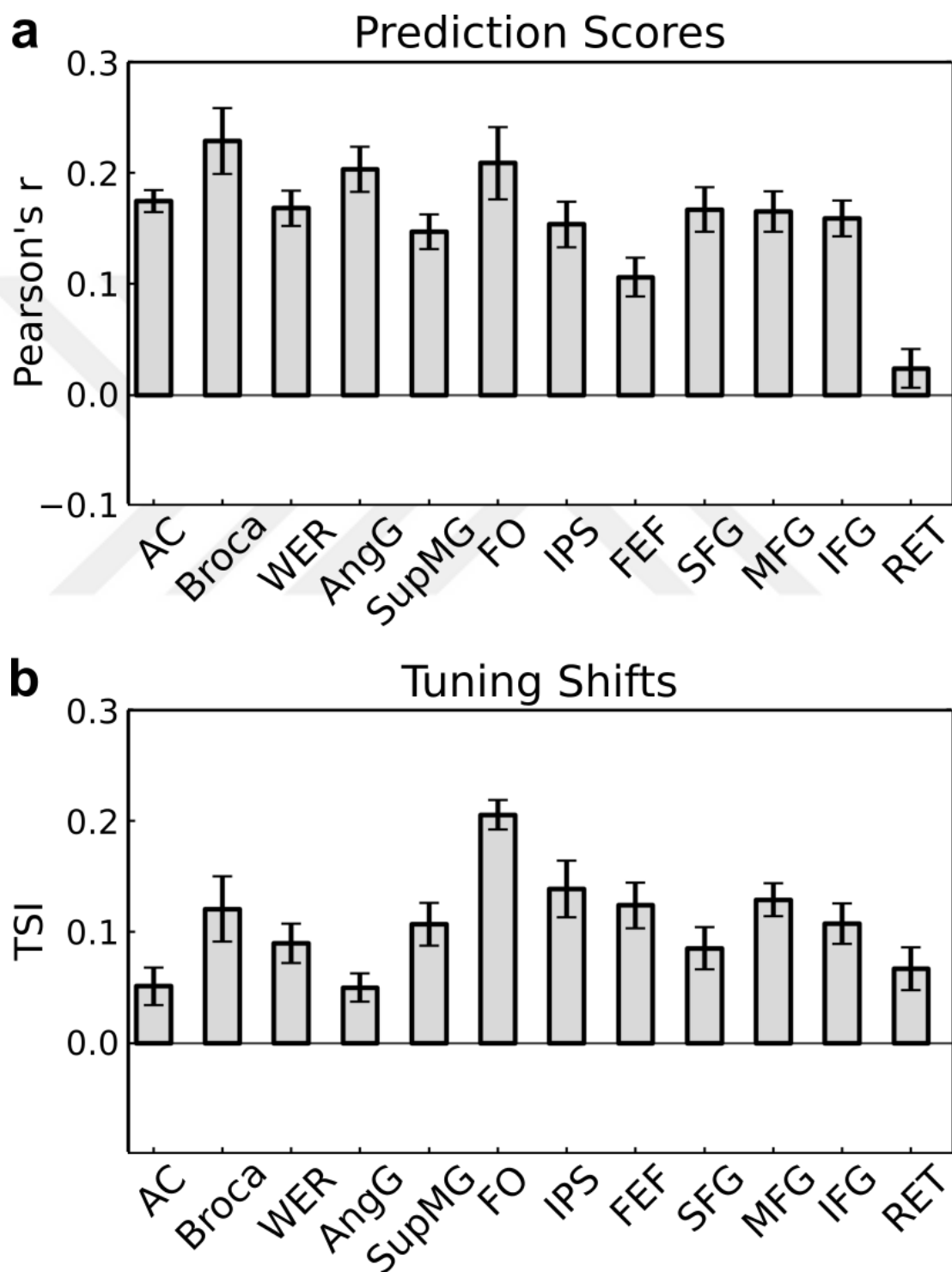


Figure 2.6: Caption next page.



Figure 2.6: **Tuning shifts in functional ROIs.** (a) Prediction scores are reported in functional ROIs as mean  $\pm$  s.e.m. across subjects. The average prediction score is  $18.65 \pm 1.75\%$  in language comprehension areas (Broca, SupMG, AngG, WER), and  $15.72 \pm 2.50\%$  in attention control areas (FO, FEF) (See 2.2.4 for ROI abbreviations). (b) Tuning shift indices are reported as mean  $\pm$  s.e.m. across subjects. Only significantly predicted voxels are considered ( $t$  test,  $q(\text{FDR}) < 0.05$ ) Attention causes varying levels of semantic tuning shifts in different functional ROIs. TSI is small in early auditory areas (AC), but larger in brain areas that correspond to later stages of language processing (WER and SupMG). The largest TSI values are seen in frontal areas that take part in language comprehension (Broca) and attention control (FO, FEF, IPS). This result implies that attentional tuning shifts become progressively stronger toward later stages of language processing and attention control.

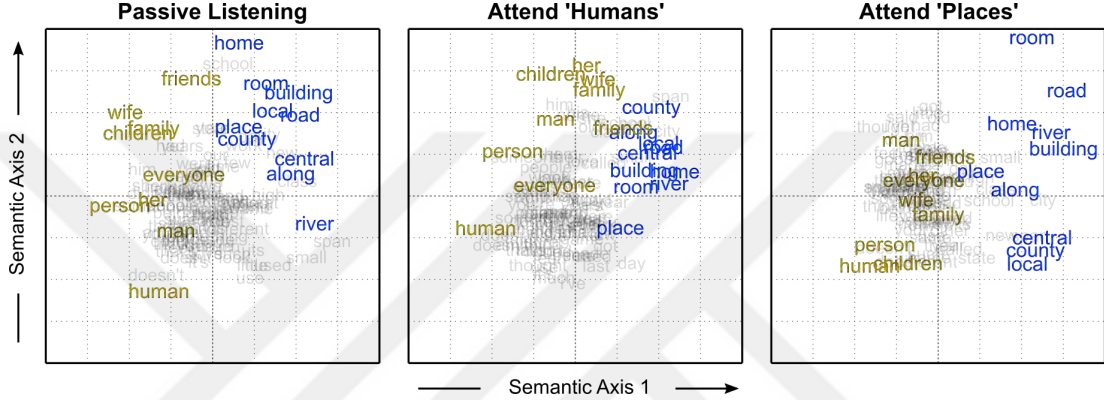


**Figure 2.7: Topology of the semantic space under passive listening and search tasks.** The tuning-shift hypothesis predicts that attention to a specific word category expands the semantic representation of target words while compressing the representation of the non-target words. To test this prediction, we estimated the topology of the semantic space independently under passive listening and the two attention tasks. PCA was performed separately on voxelwise semantic tuning vectors estimated during each task. The coefficients of a set of basic words in the first two PCs were used to visualize the topology of the semantic space. During attention tasks, words that are semantically related to the target category expand their representation, and words dissimilar to the target category are compressed. To quantify the degree of attentional warp in the semantic space, we computed a semantic warping index (see Methods). A warping index greater than 1 indicates that target words are more broadly spread in the semantic space compared to non-target words. A warping index less than 1 indicates that target words are more compactly spread in the semantic space compared to non-target words. We find that the semantic warping index is significantly greater than 1 consistently across subjects (Wilcoxon signed-rank test,  $P < 0.05$ ). These findings indicate that tuning shifts in cortex warp semantic representation in favor of attended words while compressing the representation of irrelevant words.

## Chapter 3

# Informed Feature Regularization in Voxelwise Modeling for Naturalistic fMRI Experiments

### 3.1 Introduction

Voxelwise modeling (VM) is a powerful framework to model single-voxel functional selectivity for the multitude of stimulus features that exist in natural stimuli [54, 73]. Previous studies have used this approach to sensitively model a broad range of representations from low-level spatio-temporal features to high-level object and action categories [53, 1, 74, 75]. In VM, an encoding model is constructed that contains stimulus features hypothesized to elicit BOLD responses in cortical voxels. For each voxel, a weighted linear combination of model features that best explain measured BOLD responses are computed via regularized regression. To avoid overfitting and ensure generalizability to new data, it is common to perform L2-norm regularization over model features with a uniform prior to penalize large model weights [76]. The regularization parameter is separately selected for each voxel based on a cross-validation procedure. Following parameter selection, the resulting model weights characterize the functional selectivity of single voxels.

Classical VM performs independent modeling to maximize sensitivity to single voxels, yet it disregards potential response correlations among neighboring voxels across cortex [77, 78] as well as potential stimulus correlations among model features [79]. This in turn might reduce the sensitivity of VM in capturing functional selectivity in the presence of high-levels of measurement noise. To better utilize response correlations, a spatial regularization term is proposed for VM via a graph Laplacian approach where neighborhood was defined in volumetric brain space [80]. Spatially-regularized VM models were found to improve model performance broadly across cerebral cortex. Yet, the proposed approach still uses L2-norm regularization over model features –a uniform Gaussian prior– that reflects the assumption that model features are uncorrelated [81]. This assumption might be suboptimal for models of natural stimuli. For example, in a category model that contains distinct object and action categories within natural scenes, the stimulus time courses for correlated categories are correlated (e.g., ‘human’-‘hand’, ‘kid’-‘run’, ‘building’-‘road’, ‘park’-‘tree’) [1]. Moreover, a voxel selective to humans will typically yield elevated responses to body parts, animate categories or human-related tools compared to unrelated categories [1]. Since voxelwise functional selectivity profiles for related categories are expected to be similar, model weights are not truly uncorrelated [79].

In this study, we propose a new VM approach to increase sensitivity in modeling functional selectivity (Fig. 3.1). First, we introduce an informed feature regularization that takes into account correlations among model features. This is accomplished by enforcing similarity of model weights among features that are correlated to each other. To further boost model performance, we introduce a new spatial regularization term. We had previously proposed to define voxel neighborhoods based on Euclidean distances in volumetric brain space [80]. Yet, distances along the cortical surface are more likely to capture functional similarities among neighboring voxels [82, 83]. Therefore, here we select voxel neighborhoods using geodesic distances on the cortical surface. To achieve this goal, we measured inter-voxel distances using inflated cortical space [63].

To independently evaluate improvements enabled by informed feature regularization and spatial regularization, we implemented three variants of VM: a

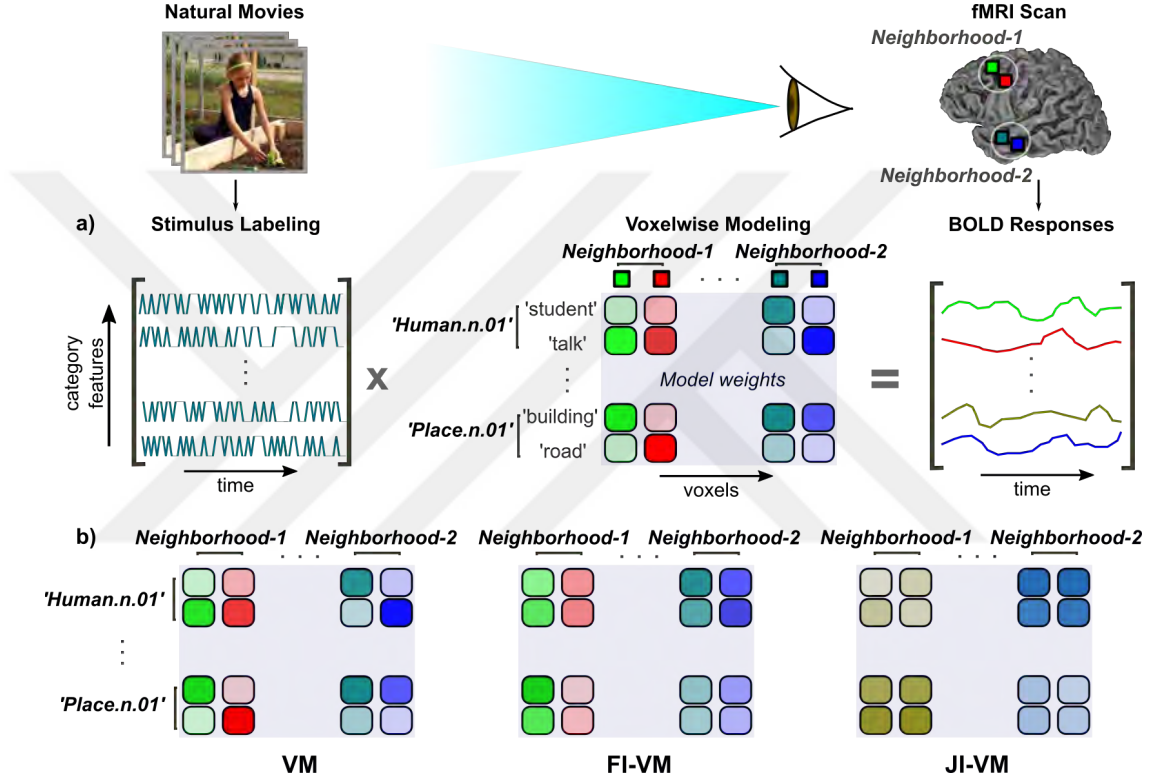


Figure 3.1: **Natural movie experiment and model fitting.** (a) Whole-brain BOLD responses were acquired while subjects viewed natural movies. To estimate functional selectivity in single voxels, we fit voxelwise models that optimally predict measured BOLD responses in terms of the category features present in the movie stimuli [1]. The resulting models describe how each of 1705 object and action categories in the movie stimulus evoke BOLD responses. (b) Classical VM ignores potential correlations in stimulus features and correlations among neighboring voxels. FI-VM takes into account correlation among model features to increase accuracy in predicting single voxel selectivity. The resulting models have increased feature coherence. JI-VM further incorporates shared information between neighboring voxels. The resulting voxelwise models have both increased feature and spatial coherence.

method that only uses informed feature regularization (Feature informed voxelwise modeling, FI-VM), a method that only uses spatial regularization (Spatially informed voxelwise modeling, SI-VM), and a method that simultaneously uses feature and spatial regularization (Jointly informed voxelwise modeling, JI-VM). Demonstrations were performed on a category model fit to whole-brain BOLD responses recorded while subjects watched a natural movie stimulus [1]. To measure correlations among 1705 object and action category features in the model, we evaluated the pairwise similarities of the categories in a word embedding space obtained from a large text corpus. These similarities were then input to a graph Laplacian term, to form the feature and spatial regularization terms. The VM variants were compared in terms of their prediction scores, and spatial and feature coherence of the resulting model weights.

## 3.2 Methods

### 3.2.1 MRI Protocols

MRI data were previously collected by researchers from The Gallant Lab on a 3T Siemens Tim Trio scanner at the University of California, Berkeley using a 32-channel Siemens volume coil. The MRI data were used in a previous study [1, 53]. Functional scans were collected using a gradient EPI sequence was used with repetition time = 2.00 s, echo time = 31 ms, flip angle =  $70^\circ$ , voxel size =  $2.24 \times 2.24 \times 4.1$  mm<sup>3</sup>, slice thickness = 3.5 mm with 18% slice gap, matrix size =  $100 \times 100$ , and field of view =  $224 \times 224$  mm<sup>2</sup> and 32 axial slices. Anatomical data for subjects for three subjects were collected using a T1-weighted multi-echo MP-RAGE sequence on the same 3T scanner. Anatomical data for the other two subjects were collected on a 1.5T Philips Eclipse scanner.

### 3.2.2 Subjects

Functional data were collected from five healthy male subjects: S1 (age 25), S2 (age 25), S3 (age 25), S4 (age 32), and S5 (age 29). All subjects had normal or corrected-to-normal vision. The experimental protocol was approved by the Committee for the Protection of Human Subjects at the University of California, Berkeley. Prior to scanning, written informed consent was obtained from all subjects.

### 3.2.3 Experimental Procedure

7200 s of continuous color natural movies were presented to the subjects in 12 separate 10 min scans ( $24^\circ \times 24^\circ$ ,  $512 \times 512$  pixels). A separate validation data were collected in 9 separate 10 min scans with same movie presented in each scan. Each validation scan contained ten separate 1 min blocks. The acquired BOLD responses were averaged across nine repetition of blocks. To minimize the effects of hemodynamic transients, data obtained in first 10 s of each run were omitted. Thus we had 3600 samples for model estimation and 270 samples for testing. Note that these same data were analyzed in several previous studies [80, 53, 84, 1]. A fixation spot ( $0.16^\circ$  square) was superimposed on the center of the screen and its color was alternated at 3 Hz, to make it continuously visible.

### 3.2.4 ROI Abbreviations

ROIs used in our study are: Auditory cortex (AC), frontal operculum (FO), intraparietal sulcus (IPS), frontal eye fields (FEF), retinotopic area (RET), fusiform face area (FFA), occipital face area (OFA), extrastriate body area (EBA), parahippocampal place area (PPA), retrosplenial cortex (RSC), occipital place area (OPA), supplementary eye fields (SEF), primary motor and somatosensory areas (S1F, M1F, S1H, M1H, S1M, and M1M), supplementary motor hand and foot areas (SMHA, SMFA).

### 3.2.5 Functional Localizers

For each subject, we defined boundaries of common functional ROIs by using functional localizers: a visual category localizer, an auditory cortex localizer, a retinotopic localizer, and a motor localizer.

In the visual category localizer, during six 4.5-minute scans the subjects were shown either places, faces, human body parts, animals, household objects, or spatially scrambled household objects. Fusiform face area (FFA) and occipital face area (OFA) were defined using the faces versus objects contrast [55]. Extrastriate body area (EBA) was defined using the human body parts versus objects contrast [56]. Parahippocampal place area (PPA), retrosplenial cortex (RSC) and occipital place area (OPA) were defined using the scenes versus objects contrast [57, 58].

For the auditory cortex localization, each subject listened to 10 repeats of a 1-minute auditory stimulus consisting of music, speech, and natural sounds. To define the auditory cortex (AC), F-statistic value across 10 repeats of each voxel’s response was used.

Retinotopic mapping data were collected in four 9-minute scans. The subjects were shown clockwise and counterclockwise rotating wedges, expanding and contracting rings in separate scans. V1, V2, V3, V4, LO, V3A, and V3B were defined based on visual angle and eccentricity maps. Using V1, V2, and V3, a broad retinotopic area (RET) was defined.

Motor localizer data were collected in a single 10-minute scan. Each subject performed small finger movements (‘hand’ cue), small toe movements (‘foot’ cue), small mouth movements (‘mouth’ cue), self-generated speech (‘speak’ cue), or saccade movements (‘saccade’ cue) in random order. Using the resulting localization data, frontal eye fields (FEF), intraparietal sulcus (IPS), frontal operculum (FO), supplementary eye fields (SEF), primary motor and somatosensory areas (S1F, M1F, S1H, M1H, S1M, and M1M) and supplementary motor hand and foot areas (SMHA, SMFA) were defined.

### 3.2.6 Data Pre-processing

Motion correction and image realignment were done using the Motion Correction FMRIB Linear Image Registration Tool (MCFLIRT) in FSL [59]. Functional images of each subject were registered to a pre-selected image of that subject. The final time series data were manually checked for accuracy. The resulting time courses were z-scored individually for each voxel.

Detrending was used to remove low frequency drifts in BOLD responses. The drifts were estimated using a median filter with a 120-second window, and then subtracted from measured BOLD responses.

Brain Extraction Tool in FSL was used to eliminate non-cortical voxels. The set of voxels within a 4-mm radius of the cortical surface were defined as cortical voxels. Following analyses were done on total of  $35,158 \pm 887$  cortical voxels (mean  $\pm$  s.d. across subjects).

### 3.2.7 Cortical Flat Maps

Cortical flat maps of each subject were generated using Freesurfer [62]. These flat maps are based on  $T_1$ -weighted anatomical scans. Anatomical surface segmentations were manually checked and corrected using pycortex [63]. Relaxation cuts were made into the surface of each hemisphere and the surface crossing the corpus callosum was removed. The calcarine sulcus cut was made at the horizontal meridian in V1. Functional images were aligned to the cortical flat maps using boundary based registration (BBR) in FSL. For visualization of flat maps, the mid-cortical surface that lies halfway between the pial and white matter surfaces was selected.

### 3.2.8 Category Model

To extract category information present in the natural movie stimulus, we used a category model that was previously introduced in another study [53, 1]. Object and action categories in each 1 s portion of the natural movies were tagged using the WordNet lexicon [85]. Finally, a list of features including 1705 distinct object and action categories was formed. The time courses of the features were resampled using a 3-lobe Lanczos filter with a cut-off frequency of 0.249 Hz (the Nyquist frequency of the fMRI acquisition). To reduce dimensionality and improve model fits, we applied PCA onto the resulting stimulus matrix. We selected only the first 300 PCs that explain 89.2% of the variance in the stimulus. PCA decorrelates model features by transforming them into a lower-dimensional subspace spanned by an orthogonal basis. Note, however, that this basis is constructed on the finite movie stimulus that only captures a relatively small set of intrinsic feature similarities. In FI-VM, these similarities are calculated based on a much broader corpus of text (The text corpus contains words from a total of 604 popular books, 2.4 million online pages and 36 million user comments scraped from reddit.com.). To rule out possible biases due to correlations between category model and global motion-energy, we added a nuisance regressor that reflects the total motion energy in the movie stimulus. To obtain the total motion energy, we summed the output of all spatiotemporal Gabor filters used in a separate motion-energy model [80].

### 3.2.9 Voxelwise Model Estimation and Validation

We used a voxelwise modeling framework to fit between stimulus and BOLD responses [53, 1]. In the classical VM approach,  $L_2$ -regularized linear regression is used to predict how each feature modulates measured BOLD responses. A category model was fit to measure voxelwise selectivity for high-level object and action categories [1]. To account for hemodynamic delays, finite-impulse-response (FIR) filter with four time delays (4, 6, and 8 s) was applied to each model feature before model fit. For BOLD responses in each time instant (TR) we used feature set of the movie stimulus at 2 TR, 3 TR, and 4 TR before the time instant.

Thus, for each TR, we have stimulus feature vector of size  $3 \times 1705$  ( $3 \times \text{Number of Features}$ ).

$$X \times W = Y \quad (3.1)$$

where  $Y$  is the response matrix of size (number of TRs  $\times$  number of voxels),  $X$  is stimulus matrix of size (number of TRs  $\times$  ( $3 \times$  number of features)),  $W$  is the weight matrix of size (( $3 \times$  number of features)  $\times$  number of voxels). In VM,  $L_2$ -regularized regression is used to minimize the cost function:

$$\min_{w_i} \sum_{i=1}^{N_{vox}} \|Xw_i - y_i\|_2^2 + \lambda \sum_{i=1}^{N_{vox}} \|w_i\|_2^2 \quad (3.2)$$

$$i = 1, \dots, N_{vox}$$

where  $w_i$  is the vector containing category model weights for voxel- $i$  and  $y_i$  is the vector containing time course of BOLD responses of voxel- $i$ . Cost function can be expressed in matrix form as:

$$\min_W \text{Tr}[(XW - Y)^T(XW - Y)] + \lambda \text{Tr}(W^T W)$$

$$\min_W \text{Tr}(W^T X^T X W) - 2 \text{Tr}(Y^T X W) + \text{Tr}(Y^T Y) + \lambda \text{Tr}(W^T W) \quad (3.3)$$

$$i = 1, \dots, N_{vox}$$

Note that minimization is still done for each voxel separately. By setting the gradient with respect to weight matrix ( $W$ ) to zero, we have:

$$2X^T X W - 2X^T Y + 2\lambda W = 0$$

$$X^T X W + \lambda W = X^T Y$$

$$(X^T X + \lambda I)W = X^T Y$$

$$W = (X^T + \lambda I)^\dagger X^T Y \quad (3.4)$$

To obtain model weights and to measure prediction scores, a 100-fold cross-validation was used. In each fold, the total of 3600 samples were split into a contiguous block of 3000 samples for model fitting and a remaining contiguous block of 600 test samples for selection of  $L_2$ -regularization parameter ( $\lambda$ ). Separate models were fit for 30 different regularization parameters in the range  $10^{-2}$  to  $10^7$  (spaced logarithmically). We calculated prediction scores based on the

coefficient of determination (i.e., explained variance,  $R^2$ ) between measured and predicted BOLD responses. Separately for each voxel, we selected the regularization parameter that maximized the average prediction score across folds. Final voxelwise models based on the optimal parameters were refit using entire 3600 time samples.

To measure final prediction scores, a 1000-fold jackknife resampling at a rate of 80% was used on a separate validation data (270 samples). For each voxel separately, final prediction score was taken as the average prediction score across jackknife folds. For easy interpretation, we multiplied prediction scores by 100, scaling them to interval of  $[0, 100]$ . To test for differences in prediction score, feature coherence or spatial coherence, the test statistic for each method was randomly sampled with replacement across subjects, and the mean difference between the methods was computed. Significance level was taken as the fraction of bootstrap samples out of 10,000 where the mean difference is less than 0 (for right-sided tests) or greater than 0 (for left-sided tests).

### 3.2.10 Parameter Optimization

In the proposed approach, there are five hyperparameters to be optimized. Three of these parameters ( $\lambda$ ,  $\lambda_f$ , and  $\lambda_s$ ) are set separately for individual voxels in each method. The remaining two hyperparameters (the standard deviation of Gaussian Filter and the feature similarity threshold) determine the Feature Laplacian matrix ( $F$ ). The Feature Laplacian matrix  $F$  is of size  $N_{\text{feat}} \times N_{\text{feat}}$ , where  $N_{\text{feat}}$  is the number of category features.

$$F = T - S \quad (3.5)$$

where  $s_{jk}$  (entries of matrix  $S$ ) is the relatedness metric of feature $_j$  and feature $_k$  which is calculated in the word embedding space (high for related features, low or zero for distant features), and  $T$  is a diagonal matrix with

$$T_{jj} = \sum_k s_{jk} \quad (3.6)$$

$s_{jk}$  indicates the cost of having feature- $j$  and feature- $k$  more apart from each other.  $s_{jk}$  is calculated within a word embedding space [86, 87, 88, 89, 90]. The word embedding space was formed based on word co-occurrence statistics. The assumption is that if two object or action categories are semantically related, then they should also be correlated in natural scenes. The co-occurrence statistics were taken from a large corpus of text. This corpus contained transcripts of popular books, Wikipedia pages, and user comments gathered from reddit.com. Co-occurrences between the words from a lexicon of 10,470 words and the 985 basis words (from Wikipedia’s List of 1000 Basic Words) were calculated. The appearance of a lexicon word within 15-word neighborhood of the basis word was taken as a single co-occurrence. We reduced the word embedding space by projecting onto the PC space obtained in 3.2.8. As a result, each of 1705 category features had a 300-dimensional vector representing its location in the reduced word embedding space. We then computed similarity (cosine similarity,  $cs_{jk}$ ) of each feature pair. We eliminated feature pairs that have a similarity smaller than the feature similarity threshold ( $d_{sim}$ ). Later, we fed those similarity values into a Gaussian filter, to get the final form of relatedness metric ( $s_{jk}$ ):

$$s_{jk} = \begin{cases} e^{\frac{-(1-cs_{jk})^2}{2\sigma^2}} & \text{if } cs_{jk} \geq d_{sim} \\ 0 & \text{if } cs_{jk} < d_{sim} \end{cases} \quad (3.7)$$

Standard deviation ( $\sigma$ ) and feature similarity threshold ( $d_{sim}$ ) were optimized a priori via a cross-validation procedure on the training data. To do this, we computed voxelwise prediction scores of separate JI-VM models with standard deviations of [0.1, 0.2, 0.4, 0.8, 1.2, 2.0] and feature similarity thresholds of [0.1, 0.2, 0.3, 0.5]. The model with standard deviation ( $\sigma$ ) of 0.4 and feature similarity threshold ( $d_{sim}$ ) of 0.2 outperformed others in most ROIs, thus we fixed these parameters throughout our study.

To incorporate spatial regularization into VM, we constructed a spatial Laplacian matrix that represents correlated information across neighboring voxels. The spatial Laplacian matrix ( $L$ ) is of size  $N_{vox} \times N_{vox}$ , where  $N_{vox}$  is the number of voxels.

$$L = P - C \quad (3.8)$$

where  $c_{il}$  (entries of matrix  $C$ ) corresponds to the proximity of voxel- $i$  and voxel- $l$  in three-dimensional space (high for immediate neighbors, low or zero for voxels far away), and  $P$  is a diagonal matrix with

$$P_{ii} = \sum_l c_{il} \quad (3.9)$$

The weight  $c_{il}$  represents the cost of having weights of voxel- $i$  and voxel- $l$  more dissimilar to each other. To compute  $c_{il}$ , first the distance between voxel- $i$  and voxel- $l$  is calculated as the Euclidean distance between those voxels in inflated cortical space [63]. Calculated distances are then input to a Gaussian filter. We have two hyperparameters in constructing the spatial Laplacian matrix  $L$ : the proximity limit after which two voxels are treated as spatially uncorrelated and the standard deviation of the Gaussian filter. Here, we prescribed these hyperparameters reported by Çelik et al. by converting into millimeter scale to use in our approach [80].

The other three hyperparameters are regularization parameters of  $L_2$ -regularization, feature regularization and spatial regularization:  $\lambda$ ,  $\lambda_f$ , and  $\lambda_s$ , respectively.  $\lambda$  indicates the degree of  $L_2$ -regularization across model features.  $\lambda_f$  indicates the degree of penalization applied to the differences within the weights of related feature pairs.  $\lambda_s$  indicates the degree of spatial regularization across category features of neighboring voxels. Since voxels across cortex may need varying degree of regularization, we optimized these parameters separately for each voxel.

### 3.2.11 JI-VM Model Estimation and Validation

Cost function for the regularized regression in the proposed approach can be expressed as:

$$\min_{w_i} \sum_{i=1}^{N_{vox}} \|Xw_i - y_i\|_2^2 + \lambda \sum_{i=1}^{N_{vox}} \|w_i\|_2^2 + \lambda_f \sum_{i=1}^{N_{vox}} \sum_{j,k} s_{jk} (w_{ij} - w_{ik})^2 + \lambda_s \sum_{i,k} c_{il} \|w_i - w_l\|_2^2 \quad (3.10)$$

$$i = 1, \dots, N_{vox}$$

where the third term is the feature regularization term and the last term is the spatial regularization term. In the feature regularization term, the weighed sum of Euclidean distances between each feature pair is calculated for each voxel separately.  $s_{jk}$  represents the cost of having model weights for feature-j and feature-k dissimilar to each other. In the spatial regularization term, the weighed sum of Euclidean distances between weights of voxels is calculated. The weight  $c_{il}$  represents the cost of having vectors containing category model weights of voxel-i and voxel-l dissimilar to each other (see 3.2.10).

Cost function in 3.10 can be expressed in matrix form as (see also 3.2.10):

$$\begin{aligned} \min_W \text{Tr} [(XW - Y)^T (XW - Y)] + \lambda \text{Tr} (W^T W) + \lambda_f \text{Tr} (W^T F W) + \lambda_s \text{Tr} (W L W^T) \\ \min_W \text{Tr} (W^T X^T X W) - 2 \text{Tr} (Y^T X W) + \text{Tr} (Y^T Y) + \lambda \text{Tr} (W^T W) + \lambda_f \text{Tr} (W^T F W) \\ + \lambda_s \text{Tr} (W L W^T) \quad (3.11) \\ i = 1, \dots, N_{vox} \end{aligned}$$

Note that minimization is still done for each voxel separately. Then, by setting the gradient with respect to weight matrix (W) to zero, we now have:

$$\begin{aligned} 2X^T X W - 2X^T Y + 2\lambda W + 2\lambda_f F^T W + 2\lambda_s W L &= 0 \\ X^T X W + \lambda W + \lambda_f F^T W + \lambda_s W L &= X^T Y \\ (X^T X + \lambda I + \lambda_f F^T) W + \lambda_s W L &= X^T Y \quad (3.12) \end{aligned}$$

### 3.2.12 Pseudo-code for JI-VM Implementation

$$\begin{aligned} (X^T X + \lambda I + \lambda_f F^T) W + \lambda_s W L &= X^T Y \\ (K + \lambda I + \lambda_f F^T) W + W B &= M \\ A W + W B &= M \quad (3.13) \end{aligned}$$

Input: X: stimulus matrix of size (time points)  $\times$  (3  $\times$   $N_{feat}$ )

Y: response matrix of size (time points)  $\times$  ( $N_{vox}$ )

$K$ : auto-covariance matrix of size  $(3 \times N_{feat}) \times (3 \times N_{feat})$ , where  $K = X^T X$   
 $M$ : cross-covariance matrix of size  $(3 \times N_{feat}) \times (N_{vox})$ , where  $M = X^T Y$   
 $\lambda$ : regularization parameter for L2-norm regularization  
 $\lambda_f$ : regularization parameter for feature regularization  
 $\lambda_s$ : regularization parameter for spatial regularization  
 $A$ : auto-covariance matrix of size  $(3 \times N_{feat}) \times (3 \times N_{feat})$ , where  $A = (K + \lambda I + \lambda_f F)$   
 $F$ : Laplacian matrix of size  $(3 \times N_{feat}) \times (3 \times N_{feat})$  which stores similarity information between category features  
 $L$ : Laplacian matrix of size  $(N_{vox}) \times (N_{vox})$  which stores proximity information between voxels  
 $B$ :  $B = \lambda_s L$   
Output:  $W$ : model weight matrix of size  $(3 \times N_{feat}) \times (N_{vox})$   
Precompute Schur decomposition of  $L$ ,  $L = U S U^T$   
Save  $U$  and  $S_d$ , where  $S_d = \text{diag}(S)$  of size  $1 \times (N_{vox})$   
**Solve:  $(AW + WB = M)$**   
begin  
  for  $\lambda$ :  
    for  $\lambda_f$ :  
      Find eigenvalues of  $A$  and store them in  $D$   
      Set  $D_d = \text{diag}(D)$ , where  $D_d$  a column vector of size is  $(3 \times N_{feat})$   
      Set  $D_r = [D_d, D_d, \dots]$  where  $D_d$  repeats  $N_{vox}$  times such that  $D_r$  is of size  $(3 \times N_{feat}) \times (N_{vox})$   
      Find eigenvectors of  $A$  and store them in  $Q$   
      for  $\lambda_s$ :  
        Set  $S_r = [S_d; S_d; \dots]$  where  $S_d$  repeats  $(3 \times N_{feat})$  times such that  $S_r$  is of size  $(3 \times N_{feat}) \times (N_{vox})$   
         $P = 1./(D_r + \lambda_s S_r)$  where  $P$  is of size  $(3 \times N_{feat}) \times (N_{vox})$   
         $W^* = -Q(P \cdot (Q^T M U))U^T$

### 3.2.13 FI-VM Model Estimation and Validation

In FI-VM, we only include feature regularization term, therefore  $\lambda_s$  is set to 0:

$$(X^T X + \lambda I + \lambda_f F^T)W = X^T Y \quad (3.14)$$

The solution to this equation:

$$W = (X^T X + \lambda I + \lambda_f F^T)^\dagger X^T Y \quad (3.15)$$

We implemented a cross-validation procedure as in the classical VM method. This time, separate voxelwise models were fit for  $30 \times 30$  different regularization parameter pairs  $(\lambda, \lambda_f)$  in the range  $10^{-2}$  to  $10^7$  (each spaced logarithmically). We measured the prediction scores using the same jackknife resampling method.

### 3.2.14 SI-VM Model Estimation and Validation

In SI-VM, we removed feature regularization from JI-VM. Equivalently, by setting  $\lambda_f = 0$  in equation 3.12, we have:

$$(X^T X + \lambda I)W + \lambda_s W L = X^T Y \quad (3.16)$$

### 3.2.15 Pseudo-code for SI-VM Implementation

$$\begin{aligned} (X^T X + \lambda I)W + \lambda_s W L &= X^T Y \\ (K + \lambda I)W + W B &= M \\ A W + W B &= M \end{aligned} \quad (3.17)$$

Input: X: stimulus matrix of size (time points)  $\times$   $(3 \times N_{feat})$

Y: response matrix of size (time points)  $\times$   $(N_{vox})$

K: auto-covariance matrix of size  $(3 \times N_{feat}) \times (3 \times N_{feat})$ , where  $K = X^T X$

M: cross-covariance matrix of size  $(3 \times N_{feat}) \times (N_{vox})$ , where  $M = X^T Y$

$\lambda$ : regularization parameter for L2-norm regularization

$\lambda_s$ : regularization parameter for spatial regularization

A: auto-covariance matrix of size  $(3 \times N_{feat}) \times (3 \times N_{feat})$ , where  $A = (K + \lambda I)$

F: Laplacian matrix of size  $(3 \times N_{feat}) \times (3 \times N_{feat})$  which stores similarity information between category features

L: Laplacian matrix of size  $(N_{vox}) \times (N_{vox})$  which stores proximity information between voxels

B:  $B = \lambda_s L$

Output: W: model weight matrix of size  $(3 \times N_{feat}) \times (N_{vox})$

Precompute Schur decomposition of L,  $L = USU^T$

Save U and  $S_d$ , where  $S_d = \text{diag}(S)$  of size  $1 \times (N_{vox})$

**Solve: (AW + WB = M)**

begin

  for  $\lambda$ :

    Find eigenvalues of A and store them in D

    Set  $D_d = \text{diag}(D)$ , where  $D_d$  a column vector of size is  $(3 \times N_{feat})$

    Set  $D_r = [D_d, D_d, \dots]$  where  $D_d$  repeats  $N_{vox}$  times such that  $D_r$  is of size  $(3 \times N_{feat}) \times (N_{vox})$

    Find eigenvectors of A and store them in Q

  for  $\lambda_s$ :

    Set  $S_r = [S_d; S_d; \dots]$  where  $S_d$  repeats  $(3 \times N_{feat})$  times such that  $S_r$  is of size  $(3 \times N_{feat}) \times (N_{vox})$

$P = 1./(D_r + \lambda_s S_r)$  where P is of size  $(3 \times N_{feat}) \times (N_{vox})$

$W^* = -Q(P.*(Q^T MU))U^T$

### 3.2.16 Feature Coherence Analysis

We hypothesized that the human brain encodes information in a way that related features elicit similar responses in single voxels. In the VM framework, this hypothesis would manifest itself as coherent model weights across related features for individual voxels. Because VM ignores feature correlations, it can have reduced sensitivity in capturing coherence among model features. Yet, informed feature

regularization enforces increased coherence among category features. Thus, in FI-VM and in JI-VM, resulting model weights are expected to be more coherent. To test this hypothesis, we defined a voxelwise feature coherence metric. We computed the contribution of voxel- $i$  to the feature regularization cost in the regression analysis:

$$\sigma_{\text{feature}}^i = \sum_{j,k} f_{jk} (w_{ij} - w_{ik})^2 \quad (3.18)$$

where the differences between weights of correlated features (i.e., for pairs of [feature- $j$ , feature- $k$ ] where  $s_{jk} > 0$ , see equations 3.5, 3.6, 3.7 and section 3.2.10 for details) are penalized according to a graph Laplacian matrix computed in the word embedding space (see 3.2.10). Voxelwise cost values were then normalized with the maximum cost obtained across the three methods (FI-, SI-, JI-VM). The normalized cost was inverted to finally obtain the feature coherence metric.

### 3.2.17 Spatial Coherence Analysis

Human brain encodes information across spatial clusters of neural populations [91, 1, 92]. Thus, it is expected that neighboring voxels represents correlated information. Since classical VM ignores spatial correlations among neighboring voxels, it is expected to fail in capturing the coherence in neighboring voxels. Due to spatial regularization, the proposed approach should be more sensitive in capturing the spatial coherence. To test this prediction, we computed a voxelwise spatial coherence metric. The spatial coherence of voxel- $i$  was taken as the mean standard deviation among category model weights of the voxels within spatial proximity of voxel- $i$  (i.e., for any voxel- $l$  where  $c_{il} > 0$ , see 3.2.11):

$$\sigma_{\text{spatial}}^i = \sum_l c_{il} \|w_i - w_l\|_2^2 \quad (3.19)$$

We normalized those values by dividing the largest value computed across the three methods and then inverted to obtain the final form of spatial coherence metric.

### 3.2.18 Noise Ceiling

The coefficient of determination ( $R^2$ ) between predicted response and measured response can be biased downward due to high measurement noise in voxel responses [93, 94, 95]. Since validation data were recorded a finite number of repeats (9 in this study), it is likely to contain high measurement noise in addition to signal. We calculated a noise ceiling for each voxel following procedures in Hsu et al. [94]. A voxelwise correction factor was calculated by dividing the maximum possible prediction score (called the noise ceiling) by the raw prediction score of each voxel. For voxels with very high level of noise, this process leads to divergent correction factors. We limited correction factors of these voxels to the average of correction factors across remaining cortical voxels.

## 3.3 Results

### 3.3.1 Prediction Scores of VM Methods

The proposed method employs additional regularization terms to consider correlations among model features as well as response correlations among voxel neighborhoods. As such, the fit models are expected to better capture voxelwise functional selectivity than classical VM. To test this prediction and to examine the contributions of the individual regularization terms, we separately computed improvements in prediction scores of FI-VM, SI-VM, and JI-VM over VM (see 3.2.9 for calculation of prediction score, see Supplemental Fig. B.1 for prediction scores of VM). Fig. 3.2 displays the change in explained variance for FI-VM, SI-VM and JI-VM in common functional regions-of-interest (ROIs, see 3.2.4 and 3.2.6). We find that FI-VM explains significantly higher variance than VM in all examined functional ROIs ( $P < 0.05$ , Bootstrap test). The average improvement in prediction score is  $1.09 \pm 0.19$  ( $\Delta R^2$ , mean  $\pm$  s.d. across ROIs, see 3.2.9 for the definition of prediction scores) for category-selective areas (PPA, FFA, EBA, OPA, and OFA), and  $0.74 \pm 0.10$  for early visual areas (RET, V4, V7). These

results suggest that informed feature regularization particularly improves model accuracy in category-selective areas. These results can be attributed to the fact that representations in category-selective areas are mediated by object and action category features, whereas those in early-visual areas are mediated primarily by retinotopic features [96, 97, 91, 1, 73, 98, 92].

We also find that SI-VM significantly improves prediction score in all ROIs ( $P < 0.05$ , Bootstrap test). These improvements are substantial broadly across cortex ( $2.75 \pm 0.55$  for early visual areas and  $2.43 \pm 0.50$  for category-selective areas). This result implies that spatial correlations among voxel neighborhoods exists across multiple stages of visual processing, and use of spatial regularization improves model accuracy across these stages.

Since JI-VM simultaneously leverages feature and spatial regularization, we reasoned that it should yield the highest level of improvement in explained variance over VM. Indeed, we find that improvement in explained variance with JI-VM is significantly higher than those with FI-VM and SI-VM in all ROIs but V4 and RSC ( $P < 0.05$ , Bootstrap test). The average increase in prediction score is  $4.57 \pm 0.56$  for category-selective areas, and  $3.25 \pm 0.59$  for early visual areas. In many ROIs including IPS, and category-selective areas, JI-VM improves prediction score as much as the sum of individual contributions from FI-VM and SI-VM. On the other hand, in RET, V4, and RSC, additional improvement by JI-VM over SI-VM seems to be limited. These results indicate that spatial and feature regularization provide independent improvements in model performance particularly across category-selective areas.

### 3.3.2 Selectivity for Model Features

In the presence of high levels of noise in training data, fit models will poorly reflect the true functional selectivity of voxels. While VM uses L2-norm regularization over model features to alleviate over-fitting, high noise levels typically lead to excessive regularization that reduce sensitivity in estimating functional selectivity. In the proposed approach, we incorporate additional regularization terms

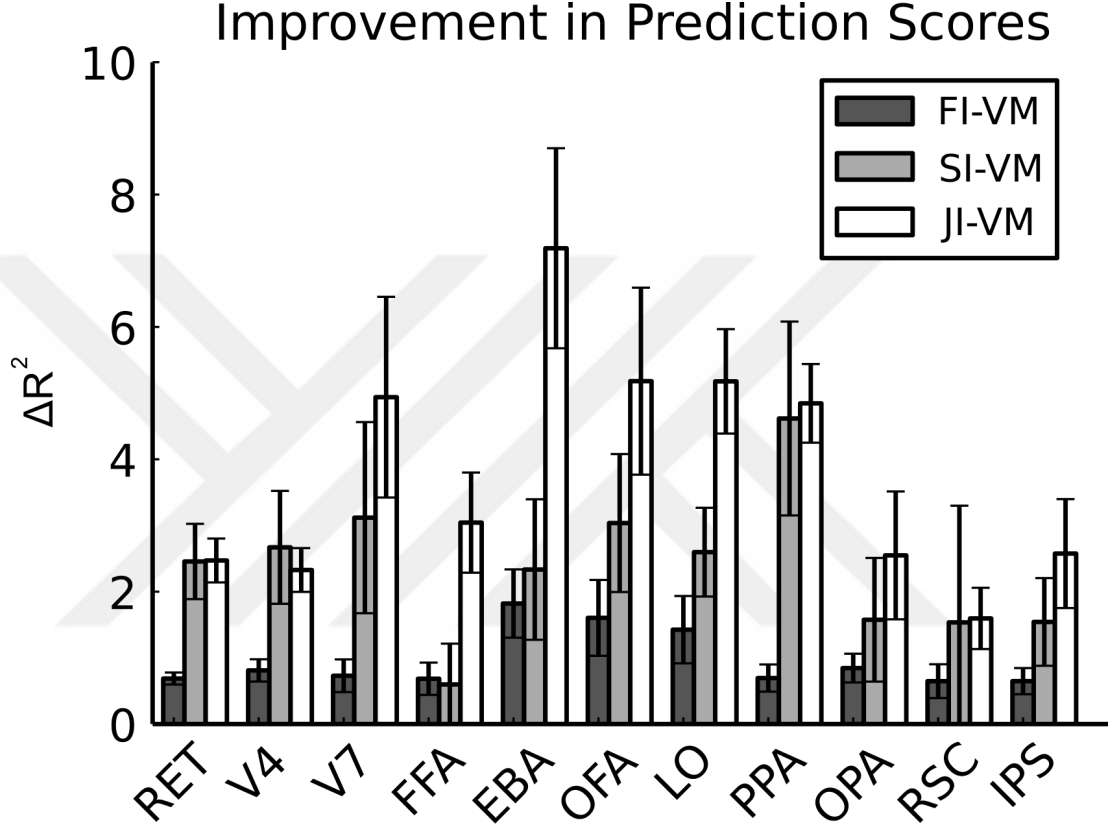


Figure 3.2: **Improvement in prediction performance over VM by FI-VM, SI-VM, and JI-VM.** The improvement in prediction performances for the three methods (FI- , SI- , JI-VM) over VM are shown for several well-known functional ROIs. Bar graphs display improvement in prediction scores ( $\Delta R^2$ , mean  $\pm$  s.d. across subjects, see 3.2.9 for the definition of prediction scores). Each of the three methods significantly improves prediction scores in all ROIs ( $P < 0.05$ , Bootstrap test). For FI-VM, improvement in prediction score is higher in category-selective areas than in early visual areas ( $1.09 \pm 0.19$  vs  $0.74 \pm 0.10$ ). For SI-VM, improvement is substantial in category-selective areas ( $2.75 \pm 0.55$ ) and in early visual areas ( $2.43 \pm 0.50$ ). JI-VM yields significantly higher prediction scores than all competing methods in all ROIs except V4 and RSC ( $P < 0.05$ , Bootstrap test). Improvement in prediction score is higher in category-selective areas than early visual areas ( $4.57 \pm 0.56$  vs  $3.25 \pm 0.59$ ). These results imply that spatial and feature regularization improve model performance particularly across category-selective areas.

across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights with a uniform Gaussian prior. To assess the level of weight penalization with each alternative method, in Fig. 3.3, we plotted the optimal  $L_2$ -regularization parameters over model features ( $\lambda$ ) on cortical flat map of a representative subject for VM, FI-VM, SI-VM and JI-VM (see Supplemental Fig. S2 for all subjects). Among the examined methods, VM applies the highest weight penalization broadly across cortex. FI-VM reduces the degree of penalization particularly in category-selective areas, as well as in early visual areas and frontal gyrus. Overall, JI-VM yields the lowest level of weight penalization consistently across cortex. These results suggest that both spatial and feature regularization decrease the amount of standard  $L_2$ -norm regularization over model features, thereby increasing sensitivity in measuring voxel selectivity to individual features.

We reasoned that informed feature regularization that incorporates feature correlations should enforce voxelwise models to have relatively similar weights on correlated category features. To examine this issue, we visualized the selectivity profiles of single cortical voxels estimated with JI-VM versus VM. Fig. 3.4 displays the selectivity profiles for JI-VM and VM for two representative voxels in EBA. In voxel-1, model weights of the subordinate categories of ‘person.n.01’ (e.g., man, woman, child...) are closer to each other with JI-VM compared to VM. Moreover, in JI-VM, many model weights are less suppressed due to decreased  $\lambda$  values. These results together indicate that informed feature regularization captures coherence in features and decreases weight penalization, thus increases sensitivity in measuring voxelwise selectivity.

### 3.3.3 Feature Coherence and Spatial Coherence of Model Weights

Classical VM ignores feature correlations so it has reduced sensitivity to capture coherence in model features. FI-VM and JI-VM, on the other hand, employ regularization that respects feature correlations, so they should better capture feature

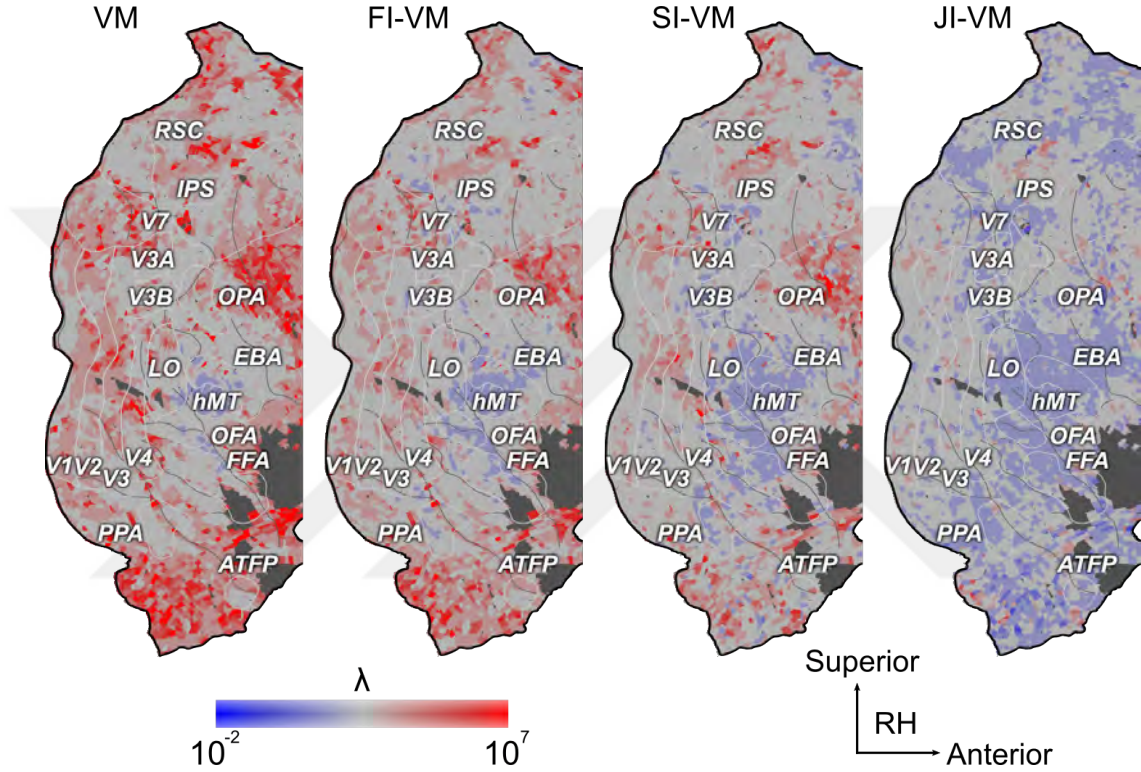


**Figure 3.3: Cortical distribution of optimal L<sub>2</sub>-norm regularization parameter.** The L<sub>2</sub>-regularization parameters over model features in four methods (VM, FI-VM, SI-VM, JI-VM) are visualized on the right hemisphere of subject S1. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, compared to VM, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity.

coherence. To test this prediction, we computed a voxelwise feature coherence metric and visualized it on flat maps (see 3.2.16). In Fig. 3.5, feature coherence of voxelwise models is shown on the posterior part of cortex of a representative subject (see Supplemental Fig. S3 for all subjects). Feature coherence of FI-VM and JI-VM models is significantly higher than VM and SI-VM in all ROIs ( $P < 0.05$ , Bootstrap test). In category-selective areas, the increase in feature coherence is substantial (see Fig. 3.6;  $10.84 \pm 0.91\%$  for FI-VM and  $17.29 \pm 1.42\%$  for JI-VM). In early visual areas and IPS, the increase is relatively modest ( $9.94 \pm 1.05\%$  for FI-VM and  $13.67 \pm 1.13\%$  for JI-VM). These results suggest that informed feature regularization increases feature coherence particularly in high-level visual areas.

Spatial regularization enforces spatially coherent model weights for neighboring voxels, so SI-VM and JI-VM should yield higher spatial coherence compared to alternative methods. To test this prediction, we computed a voxelwise spatial coherence metric and visualized it on flat maps (see 3.2.17). In Fig. 3.7, spatial coherence of voxelwise models on the posterior part of cortex of a representative subject is shown (see Supplemental Fig. S4 for all subjects). SI-VM significantly increases and FI-VM significantly decreases spatial coherence consistently in all ROIs ( $P < 0.05$ , Bootstrap test). Meanwhile, JI-VM has significantly increased spatial coherence in all ROIs ( $P < 0.05$ ), except RSC. For JI-VM, level of increase in coherence is higher in category-selective areas ( $9.37 \pm 1.40\%$ ) than in early visual areas ( $8.06 \pm 1.75\%$ ) (Fig. 3.8). These results imply that spatial regularization yields improved spatial coherence broadly across cortex. Feature coherence is increased by SI-VM in all ROIs. In spatial regularization, model weights of features are smoothed across voxel neighborhoods. We reasoned that model weights of related features may tend to get similar after spatial smoothing, causing an increase in feature coherence. Spatial coherence, on the other hand, is decreased by FI-VM in all ROIs. In FI-VM, feature regularization is done for each voxel, separately. We do not expect any increase in spatial coherence increase.

### 3.4 Discussion

Voxelwise modeling (VM) is a powerful tool to model single voxel selectivity to natural stimulus features with high sensitivity [54, 73]. Still, VM can yield suboptimal performance in the presence of high levels of measurement noise. Classical VM does not consider potential correlations among model features and spatial correlations among neighboring voxels. To improve performance, here we first introduced an informed feature regularization that considers feature correlations. We also introduced a spatial regularization approach that constructs voxel neighborhoods based on cortical distances. Our results indicate that proposed approach provides more predictive voxelwise models for a natural vision experiment in many regions across cortex. We further measured coherence of fit models and our results suggest that improvements in prediction performance can be attributed to increased spatial and feature coherence across cortical representations.

With similar motivations to treat feature correlations, a recent study calculated a linear transformation of a category model to impose regularization with a non-uniform prior on model features [79]. The transformation was defined based on the dimensions of a word embedding space constructed using co-occurrence statistics in a large text corpus. It was reported that the proposed transformation approach yielded improved prediction scores for category models across significantly-predicted cortical voxels. As opposed to a transformation of model features, here we leave the model features in their original space and leverage potential correlations through a graph Laplacian based regularization term. Avoiding an additional transformation might offer an advantage in terms of interpretability of resulting model weights. Our approach also allows for independent weighing of differences of model weights for each pair of category features. It remains important future work to comparatively demonstrate the transformation versus regularization approach for treatment of feature correlations.

Regularization across model features has also been leveraged in neuroimaging studies beyond fMRI [99, 100] or studies that focus on modeling natural

stimuli [101, 102]. A typical way that feature correlations are leveraged is to reduce dimensionality by eliminating redundant features from subsequent analyses. A previous study incorporated feature regularization to improve AD diagnosis performance with both regression and classification models [100]. The authors estimated similarity of neuroimaging features on MRI and PET images, and then used these estimates to regularize correlated features during modeling. They have reported increased regression and classification performance in AD diagnosis. In another study on medical risk stratification, the authors implemented feature regularization on various features extracted from Electronic Medical Records (EMRs) on different time scales during ordinal regression [99]. They have reported that suicide risk can be better predicted by a model of EMR-derived features compared than by clinicians. In both studies, the authors discarded redundant features via an L1-norm regularization. Unlike these previous studies, here we retained all stimulus features in the model. Several previous studies have employed a feature regularization approach similar to the one proposed here for robust text classification [101, 102]. In Zhang et al., the authors regularized the correlated subtopics for a discriminative classifier. They constructed a feature affinity matrix using subtopic co-occurrences on a large unlabeled dataset as in our method. They, later, included a feature regularization term based on the feature affinity matrix into maximum entropy training objective. They have reported that feature regularization on subtopic features improved text classification accuracy over other semi-supervised learning approaches. In Sandler et al., the authors incorporated feature regularization into sentiment classification and newsgroups document classification via logistic regression. Similar to our approach, they used a graph Laplacian based regularization by using both word co-occurrences and prior domain information to compute feature similarities. They have reported improved classification performances over many other semi-supervised learning methods. Furthermore, to utilize stimulus correlations, we may have imposed multivariate priors on model weights [103]. This process is equivalent to transforming the feature space into another one. Still, resulting feature space may not be optimal for all cortical voxels. Here, our proposed method controls the level of feature regularization through a hyperparameter to optimally regularize model weights across all voxels.

The proposed method leverages a linguistic model to guide feature regularization on a visual category model. Indeed, several recent studies suggest that category information in natural visual scenes is well captured by semantic features in natural language. Clarke et al. reported that visual objects that shared a superordinate category yielded more similar activation patterns in perirhinal cortex than those that did not [104]. Carlson et al., reported a correspondence between semantic relationships of objects presented and neural representations of these objects in inferior temporal cortex [105]. Semantic relatedness was measured using hierarchical word relations and distributional patterns of words in large text corpora. More recently, Wen et al. reported for a pair of visual objects significant correlation between representational similarity across cortex and semantic [106]. Semantic similarity was taken as distance in word embedding models [107, 108].

Here, we used a common set of hyperparameters for constructing the feature Laplacian matrix and spatial Laplacian matrix. A potential avenue for improvement for the proposed method is to consider voxelwise optimization of these parameters. Note, however, that this will require considerably higher computational load. Improving performance in feature regularization might also be viable via different approaches to computing similarity of model features. Here, we calculated them in the word embedding space constructed using co-occurrence statistics. One can typically utilize another word embedding space constructed by word2vec or LSA to find the similarities of model features [107, 109].

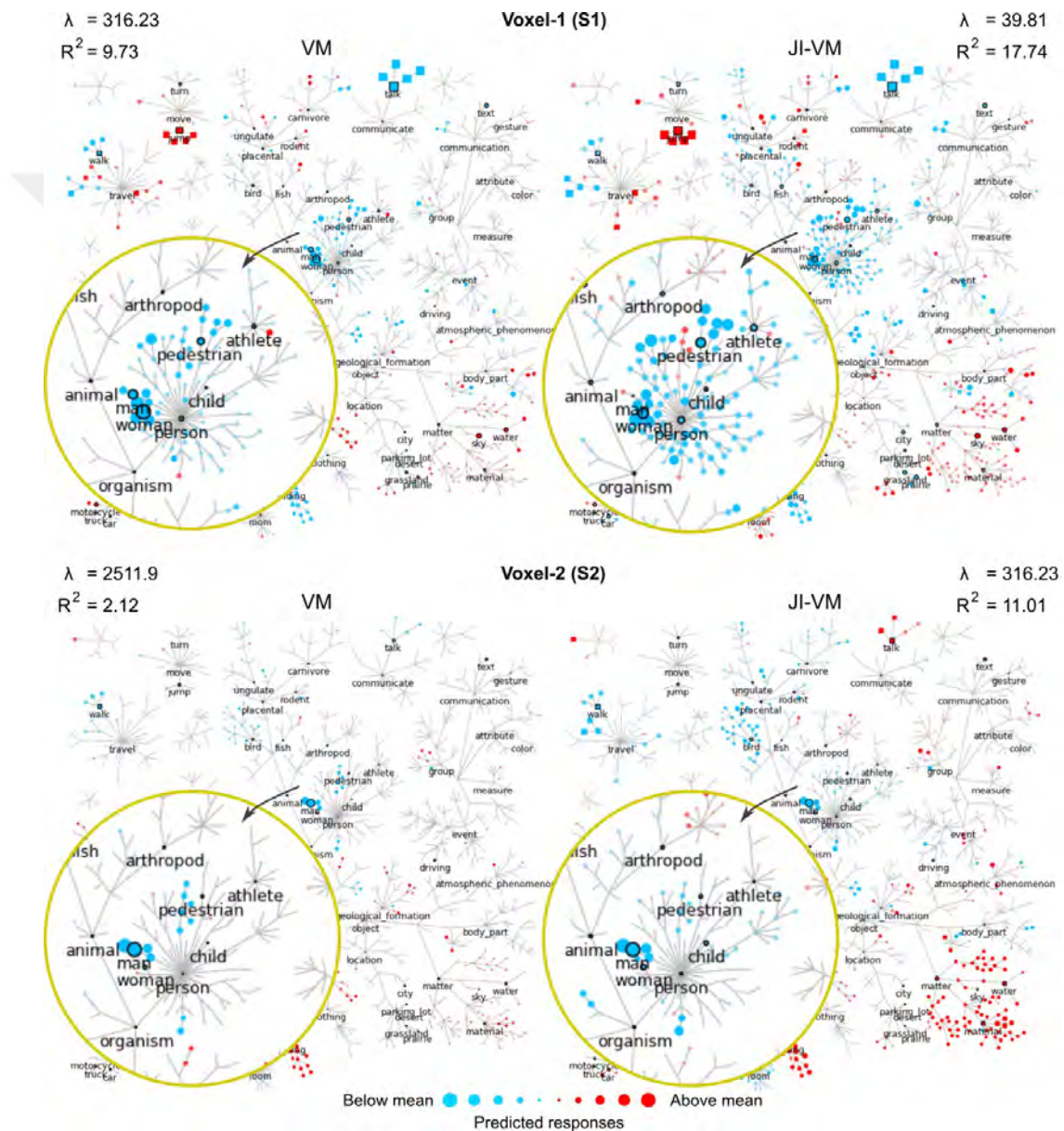


Figure 3.4: Caption next page.



Figure 3.4: **Functional selectivity in a single voxel.** Functional selectivity of two well-modeled voxels in EBA are visualized for several object and action categories organized under the WordNet hierarchy. Each node represents the estimated response of the voxel to the labeled object or action category. Red nodes correspond to categories that evoke above-mean responses and blue nodes correspond to categories that evoke below-mean responses. Sizes of the nodes indicate the magnitude of the evoked responses. In the proposed approach, we incorporate feature and spatial regularization that limit unnecessary penalization of model weights by  $L_2$ -regularization ( $\lambda = 316.23$  vs.  $\lambda = 39.81$  for voxel-1 and  $\lambda = 2511.9$  vs.  $\lambda = 39.81$  for voxel-2), while improving prediction scores ( $R^2 = 9.73$  vs.  $R^2 = 17.74$  for voxel-1 and  $R^2 = 2.12$  vs.  $R^2 = 11.01$  for voxel-2). As a result, in VM, the model weights are relatively suppressed compared to JI-VM. Furthermore, feature regularization enforces assignment of more similar weights to correlated categories in voxelwise models. For example, in voxel-1, model weights of the subordinate categories of ‘person.n.01’ (e.g., man, woman, child...) are closer to each other in JI-VM compared to VM. Taken together, these results imply increased sensitivity in capturing functional selectivity.

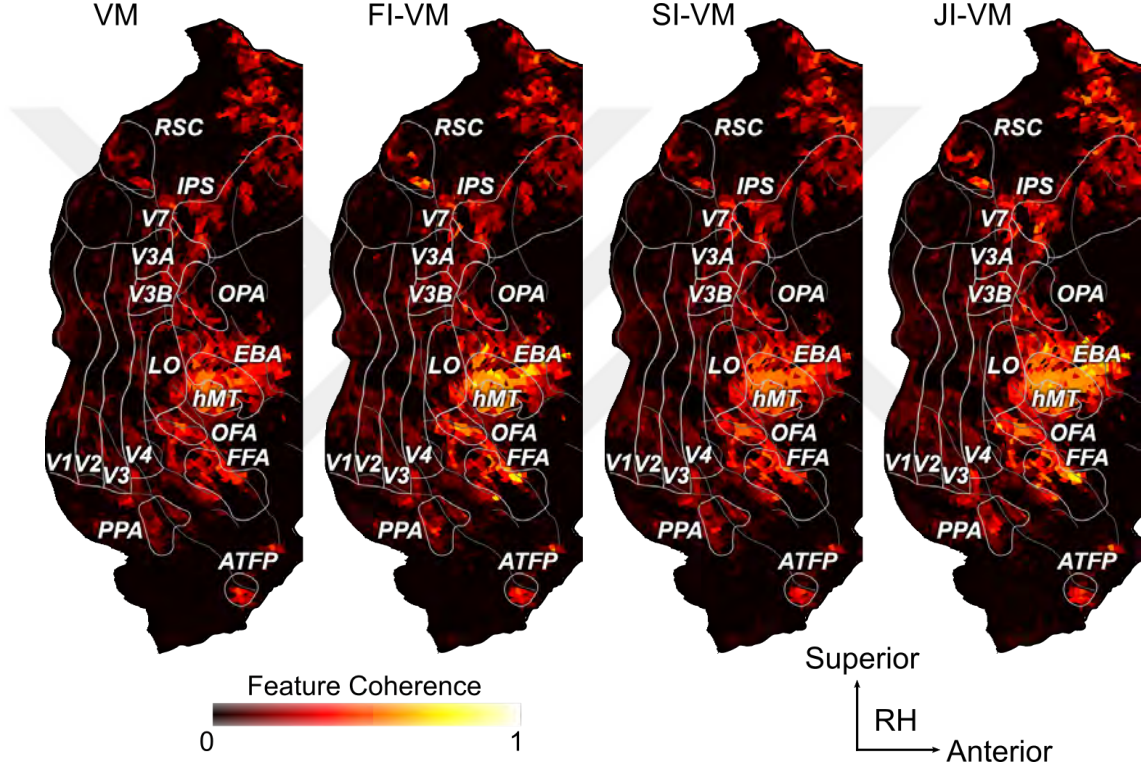


Figure 3.5: **Feature coherence for VM, FI-VM, SI-VM, and JI-VM.** Feature coherence for subject S1 is shown on the posterior part of the right hemisphere. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. Informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM is higher than that of VM and SI-VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITP. These results indicate that informed feature regularization better captures feature coherence broadly across visual cortex.

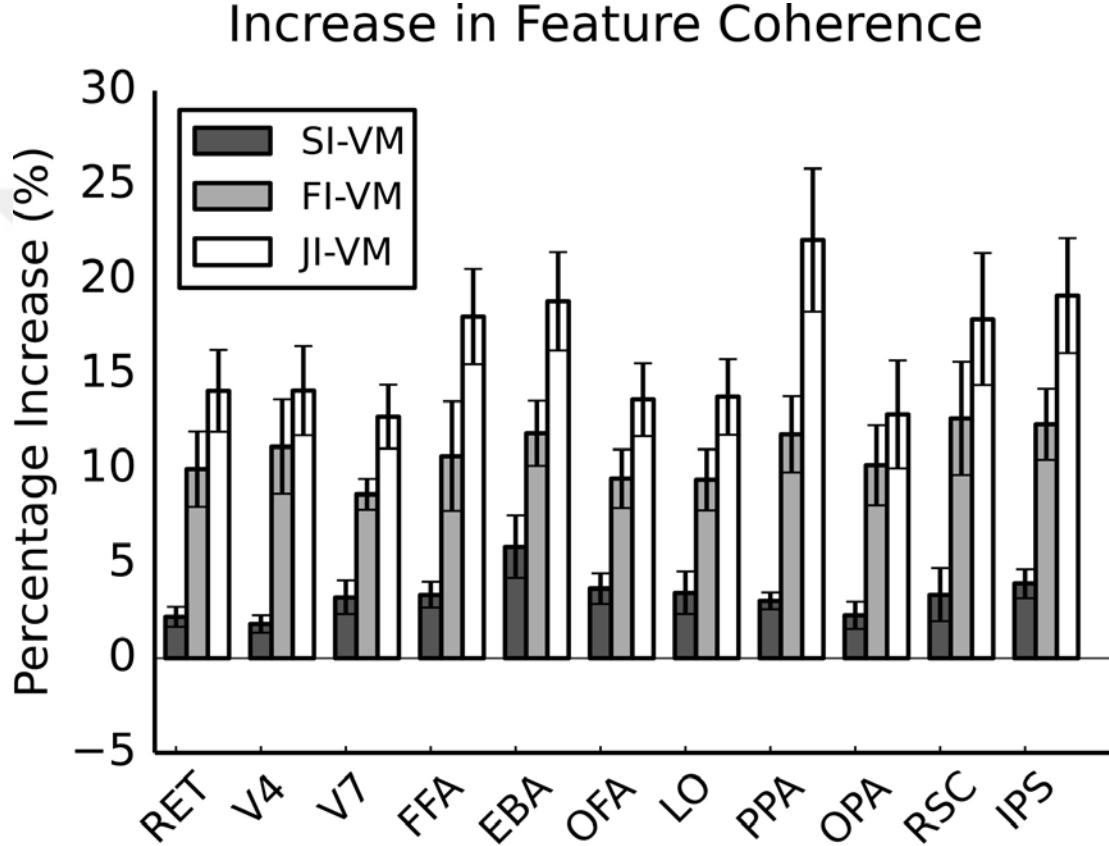


Figure 3.6: **Feature coherence across functional ROIs.** Bar graphs display increases in feature coherence by three methods (SI-VM, FI-VM, JI-VM) over VM (mean  $\pm$  std across subjects). Feature coherence of FI-VM and JI-VM models is significantly higher than that of VM in all functional ROIs ( $P < 0.05$ , Bootstrap test). The average increase in category-selective areas is  $10.84 \pm 0.91\%$  for FI-VM and  $17.29 \pm 1.42\%$  for JI-VM, whereas the average increase in early visual areas is  $9.94 \pm 1.05\%$  for FI-VM and  $13.67 \pm 1.13\%$  for JI-VM. In contrast, SI-VM that lacks feature regularization has relatively limited increase in feature coherence across all ROIs. These results suggest that informed feature regularization increases functional coherence in cortical representations particularly in high-level visual areas.

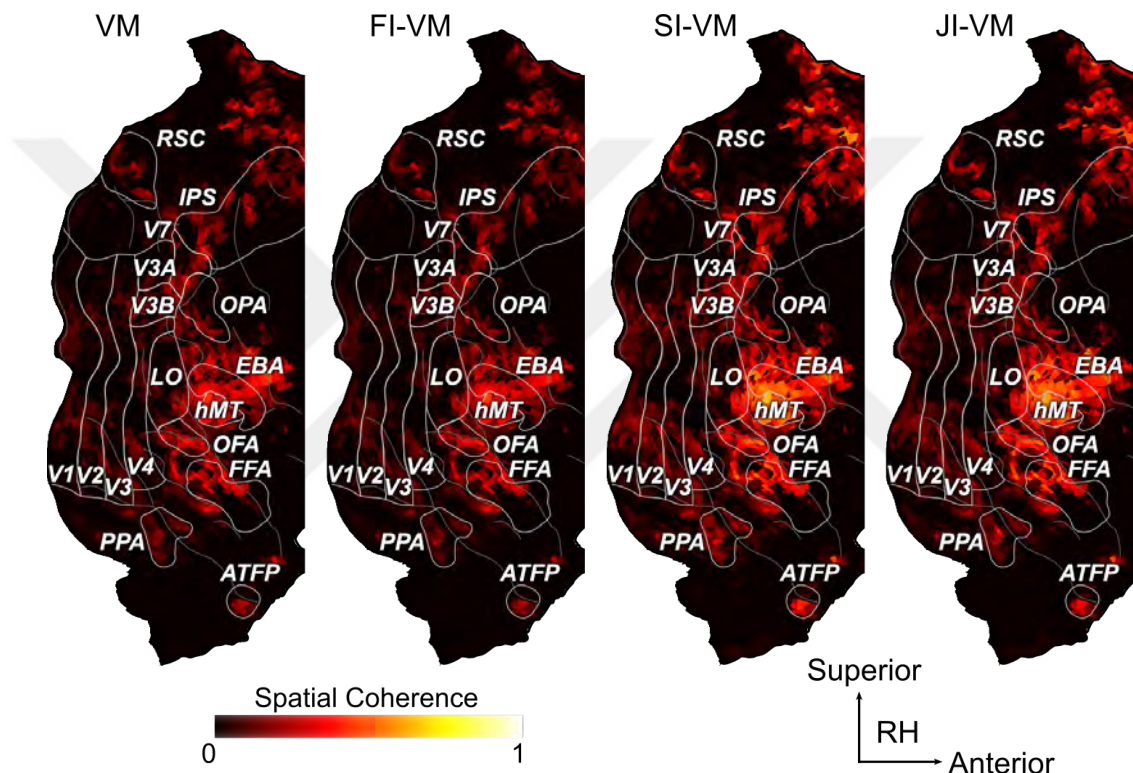


Figure 3.7: **Spatial coherence for VM, FI-VM, SI-VM and JI-VM.** Spatial coherence for subject S1 is shown on the posterior part of the right hemisphere. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, SI-VM and JI-VM yields higher spatial coherence consistently across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and ATRP. These results suggest that spatial regularization better captures spatial coherence broadly across visual cortex.

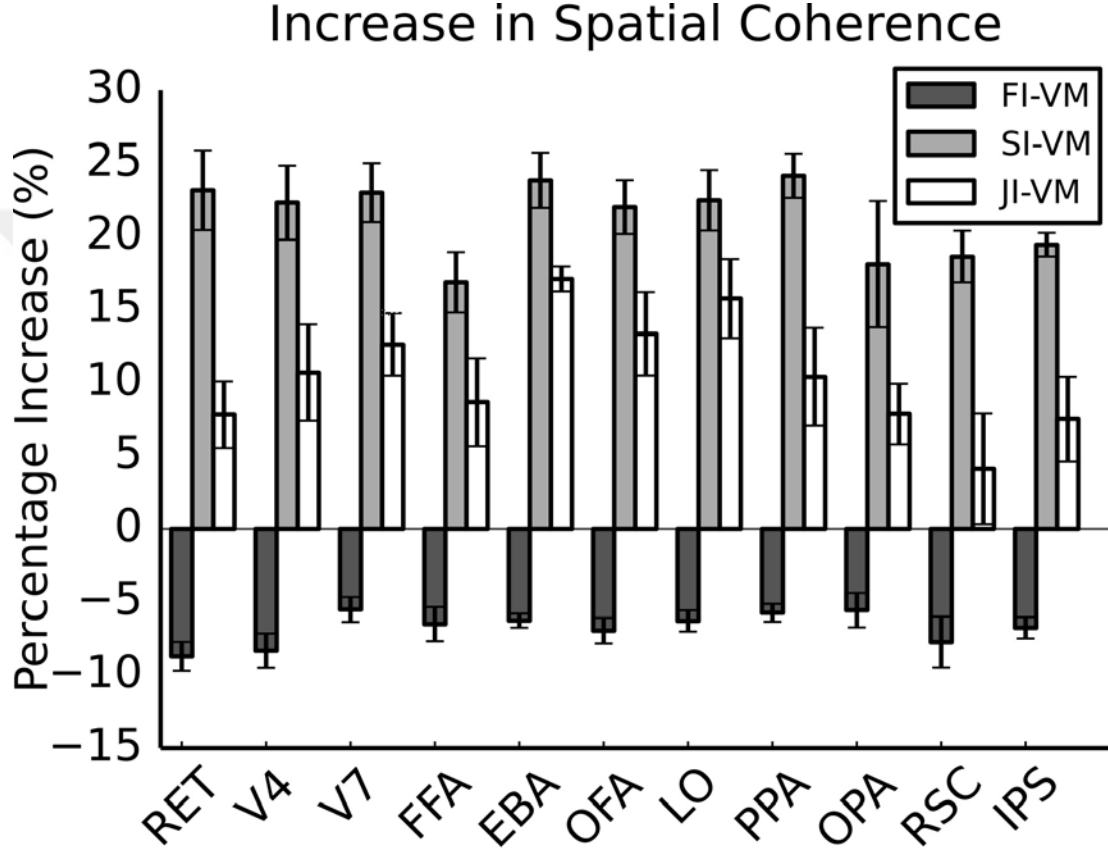


Figure 3.8: **Spatial coherence across functional ROIs.** Bar graphs display increases in spatial coherence by three methods (FI-VM, SI-VM, JI-VM) over VM (mean  $\pm$  std across subjects). Spatial regularization enforces increased spatial coherence among cortical representations of neighboring voxels. SI-VM and JI-VM have significantly higher spatial coherence than VM in all functional ROIs ( $P < 0.05$ , Bootstrap test). The average increase in category-selective areas is  $19.78 \pm 1.45\%$  for SI-VM and  $9.37 \pm 1.40\%$  for JI-VM, whereas the average increase in early visual areas is  $21.99 \pm 1.41\%$  for SI-VM and  $8.06 \pm 1.75\%$  for JI-VM. Meanwhile, FI-VM that lacks spatial regularization does not yield any improvement in spatial coherence. These results indicate that spatial regularization increases spatial coherence among cortical representations of voxel neighborhoods.

## Chapter 4

# Enhanced Phase-Sensitive SSFP Reconstruction for Fat-Water Separation in Phased-Array Acquisitions

### 4.1 Introduction

Balanced Steady-State Free Precession (bSSFP) sequences typically generate relatively higher levels of signal from fat when compared to water tissues. However, separation of these two resonances is critical for many applications including cartilage imaging [21], abdominal imaging [22], and angiography [110, 24]. Among the approaches proposed to address this problem were steady-state techniques that reshape magnetization profiles [111, 23, 112], techniques that temporarily alter transient signal profiles [113, 114, 115], and Dixon-type techniques that rely on multiple acquisitions to separate the two resonances [116, 117, 118, 119]. Because the above techniques either suppress fat signals during acquisition or perform subvoxel fat–water decomposition based on multiple signal measurements, they

can offer reduced sensitivity to partial volume effects. At the same time, however, these techniques require pulse sequence modifications and multiple echoes or acquisitions that usually prolong scan times.

Phase-sensitive SSFP (PS-SSFP) is a scan-time-efficient alternative for fat–water separation that does not require modification of standard bSSFP sequences [120]. PS-SSFP generates out-of-phase fat and water signals by placing their respective resonances in neighboring bSSFP passbands. Following a correction for additional slow-varying phase components due to inadvertent factors such as field inhomogeneity and coil sensitivity, PS-SSFP leverages abrupt phase changes arising near fat–water boundaries to separate the two signals. However, residual phase variations that remain in corrected images can cause suboptimal fat–water separation [121].

PS-SSFP reconstructions have been previously demonstrated to work reliably on images acquired with quadrature coils [120, 121]. Unlike quadrature coils with uniform coil sensitivity, individual channels in phased-array coils typically have relatively large spatial variations in sensitivity, resulting in increased noise while estimating image phase away from the coil centers. While an earlier study has proposed to address this issue by performing phase-sensitive reconstruction on a single combined image across coils [122], inter-coil inconsistencies in image phase may cause phase errors in the combination and deteriorate reconstruction performance. Other important studies have proposed performing separate phase-sensitive reconstructions on individual-coil images, which were then combined by aligning intercoil phase offsets to prevent fat–water swaps [22, 123]. However, residual phase errors between individual-coil images may show complex spatial variations and thus may not fully correctable by global factors.

Here we propose and assess an improved strategy for phase-sensitive fat–water separation in bSSFP images acquired with phased-array coils. The specific aims of this study were to improve the accuracy of region-growing phase correction via an unsupervised seed-selection procedure, to enhance tissue classification by optimally aggregating phase information across individual-coil images, and finally to demonstrate the performance of the proposed strategy in vivo on bSSFP

angiograms acquired in the lower extremities.

## 4.2 Methods

### 4.2.1 Dual-Acquisition PS-SSFP Reconstruction

The PS-SSFP approach relies on the frequency dependence of bSSFP magnetization profiles [120]. Balanced SSFP sequences generate periodic magnetization profiles, with  $\pi$ -radians phase difference between consecutive passbands (Fig. 4.1a). Therefore, a  $\Delta\phi = \pi$  phase-cycled bSSFP sequence can produce out-of-phase images when the repetition time (TR) is selected to place fat and water resonances an odd number of passbands apart,  $TR = (2n+1)/\Delta f$  where  $n$  is an integer and  $\Delta f$  is the fat–water frequency difference (eg,  $TR = 4.6$  msec at 1.5T). To improve immunity against off-resonance while maintaining this phase difference, dual-acquisition PS-SSFP [121] acquires a separate  $\Delta\phi = 0$  phase-cycled bSSFP image, and complex sums the  $\Delta\phi = \pi$  and  $\Delta\phi = 0$  images (Fig. 4.1 a). The acquired signal can be expressed as:

$$S(r) = [W(r) - F(r)] \cdot e^{-i\theta(r)} \quad (4.1)$$

where  $r$  is the spatial location,  $W$  is the level of water signal,  $F$  is the level of fat signal, and  $\Theta$  denotes the spatially varying image phase due to various sources including field inhomogeneity, susceptibility, and coil sensitivity.

PS-SSFP assumes that each voxel contains dominantly fat or water tissue, and that  $\Theta(r)$  is distinguishable from the rapid phase shifts near fat–water boundaries. In such cases, a region-growing phase-correction can be applied to remove the slowly varying phase component,  $\Theta(r)$  [22, 119, 120]. In the current study, we use the correction algorithm proposed by Hargreaves et al [121]. This algorithm first splits the imaging volume into small blocks (eg,  $6 \times 6 \times 6$  voxels). Starting with a seed block, it updates the block phases by  $\pi$ -radians when necessary, such that the intensity-weighted sum of block phases varies smoothly across the image.

Following correction of the images for the estimated block phases, phase-sensitive fat–water separation is performed.

The reliability of PS-SSFP can be compromised by rapid phase variations that violate the assumption of gradually varying  $\Theta$ . Unlike quadrature coils with broad spatial coverage, phased-array coils typically introduce large phase variations away from the coil centers where the individual-coil sensitivity diminishes. To demonstrate typical differences in image phase between quadrature-coil and phased-array data, we acquired bSSFP images of the lower leg with  $TR = 9.2$  msec at 1.5T using both a quadrature coil and a receive-only 8-channel phased-array coil. This  $TR$  selection—an even multiple of the ideal  $TR$ —places fat and water resonances two passbands apart, and thus the resulting tissue signals are in phase. The distributions of voxel magnitude and phase across in-phase bSSFP images were visualized (Fig. 4.1b). As expected, voxel phases are much more broadly spread and variable in a phased-array coil compared to a quadrature-coil. These excessive phase variations in turn increase the possibility of fat–water misclassification in phased-array bSSFP acquisitions.

#### 4.2.2 PS-SSFP for Phased-Array Acquisitions

We propose a strategy to improve phase-sensitive reconstructions of bSSFP data acquired with phased-array coils. In regular PS-SSFP, a central voxel is selected as the starting seed for region-growing phase correction [120]. A single, common seed can be designated for all coils to prevent global phase inconsistencies among coils [22]. However, in coils for which the starting seed is relatively distant to the peak sensitivity region, this procedure can cause suboptimal phase correction and fat–water swaps in subsequent reconstructions. To address this issue, we propose an unsupervised technique for coil-wise seed selection (Fig. 4.2a). This technique first reconstructs a low-resolution image for each coil by Fourier transforming the central 4% of  $k$ -space data. The coil-sensitivity profiles ( $C_{1:N}$ ) are then taken as the ratio of these individual-coil images to the sum-of-squares combination of low-resolution images across coils [124]. Note that fat typically yields much

higher bSSFP signals than water tissues in the lower extremities considered here [125]. A fat-voxel seed can thus be selected based on signal intensity. However, since tissue boundaries may be blurred in low-resolution sensitivity profiles, the sensitivity peak may correspond to a water or background voxel neighboring a high-intensity fat voxel. For this reason, the sensitivity profile is multiplied with the magnitude of the respective full-resolution coil image prior to selection. The voxel that maximizes this product is then designated as the seed for each coil. This procedure ensures that seeds are in close proximity to the sensitivity peaks.

Next, we employ a separate region-growing phase correction on each coil with the previously identified seeds to remove the slowly varying phase components (Fig. 4.2b). We initiate the phase correction in each coil on a high-signal fat voxel. As a result, fat voxels in corrected coil images will tend to cumulate around 0 radians (positive real part), whereas water voxels will cumulate around  $\pi$ -radians (negative real part) [120]. This seed-selection procedure thereby accounts for inter-coil phase offsets, avoiding global fat–water swaps across separate coils. However, correction accuracy can still degrade in regions away from the coil center where reduced coil sensitivity yields low image magnitude and noisy phase. Therefore, we propose to obtain accurate phase estimates by pooling information across coils. Here we assume that the coil elements in the phased array collectively provide sensitive coverage across the entire imaging field of view (FOV). The following linear combination of coil images can then be computed [124]:

$$S_{cmb}(r) = \frac{\sum_{i=1}^N S_i^{pc}(r) \cdot |C_i(r)|}{\sum_{i=1}^N |C_i(r)|} \quad (4.2)$$

where  $S_i^{pc}$  denotes the phase-corrected image from the  $i^{\text{th}}$  coil, and  $N$  is the number of coils. Note that the coil-sensitivity profiles estimated from central  $k$ -space data contain residual image phase. To prevent inter-coil inconsistencies due to this remnant phase, the phase-corrected coil images in Eq. 2 are weighted by the magnitude of coil sensitivities. Finally, the phase of  $S_{cmb}$  is used to classify each voxel as water or fat.

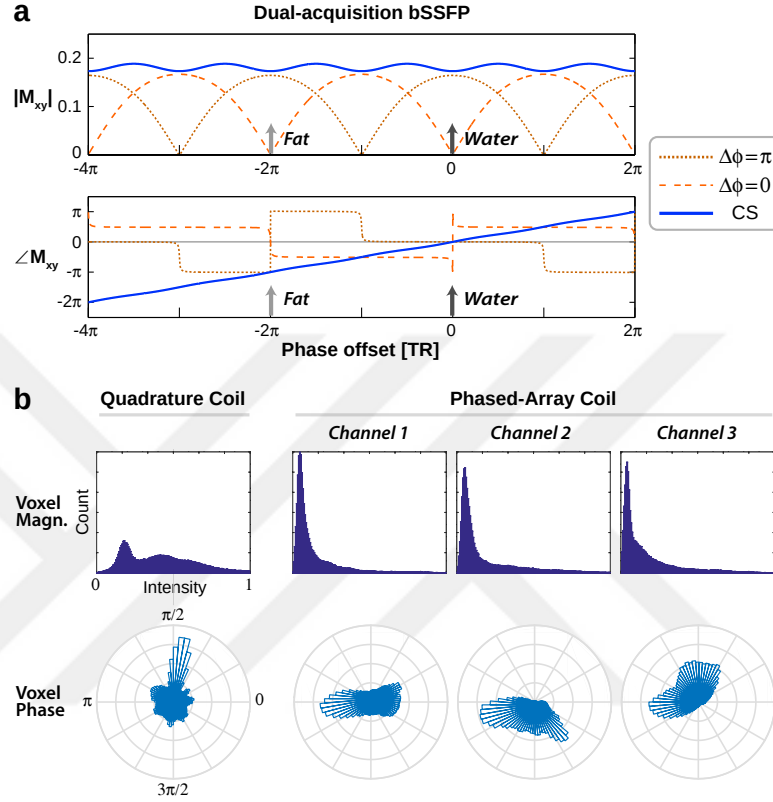


**Figure 4.1: Balanced Steady-State Free Precession.** (a) Transverse magnetization profiles of  $\Delta\phi = \pi$  (dotted line) and  $\Delta\phi = 0$  (dashed line) phase-cycled bSSFP sequences were simulated along with their complex-sum (CS, solid line). The simulations assumed  $90^\circ$  flip angle,  $TR/TE = 4.6/2.3$  msec and  $T_1/T_2 = 1200/250$  msec at 1.5T. Magnitude (top row) and phase (bottom row) profiles are shown separately, and the locations of fat and water resonances are marked with arrows. Fat and water signals are  $\pi$ -radians out of phase for  $\Delta\phi = \pi$  but large field inhomogeneity may cause water or fat resonances to leak into neighboring bands. CS maintains the  $\pi$ -radians fat–water phase difference while reducing sensitivity to field inhomogeneity. (b) To detect this phase difference, PS-SSFP uses a correction step to remove phase variations from additional sources including coil sensitivity. To illustrate the effects of coil sensitivity, in-phase bSSFP images ( $TR/TE = 9.2/4.6$  msec at 1.5T) were acquired using a quadrature extremity coil and an 8-channel phased-array coil (three representative channels are shown). Histograms of voxel magnitude (top row) and polar histograms of voxel phase (bottom row) are visualized. Because individual channels in the phased-array coil have compact spatial coverage, voxel magnitudes accumulate around lower intensities compared to the quadrature-coil image. The narrow sensitivity profiles also cause the voxel phases to be more broadly spread in phased-array images.

### 4.2.3 Alternative PS-SSFP Reconstructions

To compare the performance of the proposed multiseed combined-coil method, we implemented three alternative PS-SSFP reconstructions:

1. Single-seed individual-coil reconstruction ( $\text{PS}_{\text{ss}}$ ): In  $\text{PS}_{\text{ss}}$ , a fat voxel centrally located within the volume was selected manually as a common seed for all coils. Phase correction was performed individually on each coil image starting with this common seed. Following phase correction, global phase offsets among coil images were removed, and individual-coil water images were reconstructed [22, 123]. The final reconstruction was calculated as a weighted combination of individual-coil water images, where each coil's weight was proportional to the magnitude of its sensitivity, as in Eq. 2.
2. Multiseed individual-coil reconstruction ( $\text{PS}_{\text{ms}}$ ): In  $\text{PS}_{\text{ms}}$ , the starting seeds were determined according to the unsupervised selection procedure employed in the proposed method. The remaining parts of  $\text{PS}_{\text{ms}}$  were identical to the  $\text{PS}_{\text{ss}}$  reconstruction. Thus, water images were reconstructed separately for each coil and then linearly combined.
3. Single-seed combined-coil reconstruction ( $\text{PS}_{\text{cc}}$ ): The first step of  $\text{PS}_{\text{cc}}$  was to calculate a linear combination of coil images [124]:

$$S_{\text{cc}}(r) = \frac{\sum_{i=1}^N S_i(r) \cdot C_i^*(r)}{\sum_{i=1}^N |C_i(r)|} \quad (4.3)$$

where uncorrected coil images  $S_i$  are multiplied with the complex conjugate of the coil sensitivities  $C_i^*$ . Note that weighting uncorrected coil images by the sensitivity magnitudes as in Eq. 2 would cause signal loss due to intercoil phase inconsistencies. Instead, the combination in Eq. 3 removes from individual-coil images the low-spatial-frequency phase variations captured by the sensitivity estimates (including global phase terms). To account for gradual phase variations still remaining in the combination, region-growing phase correction was performed on  $S_{\text{cc}}$  using a central fat-voxel seed, and the resulting phase-corrected image was used for fat-water separation.

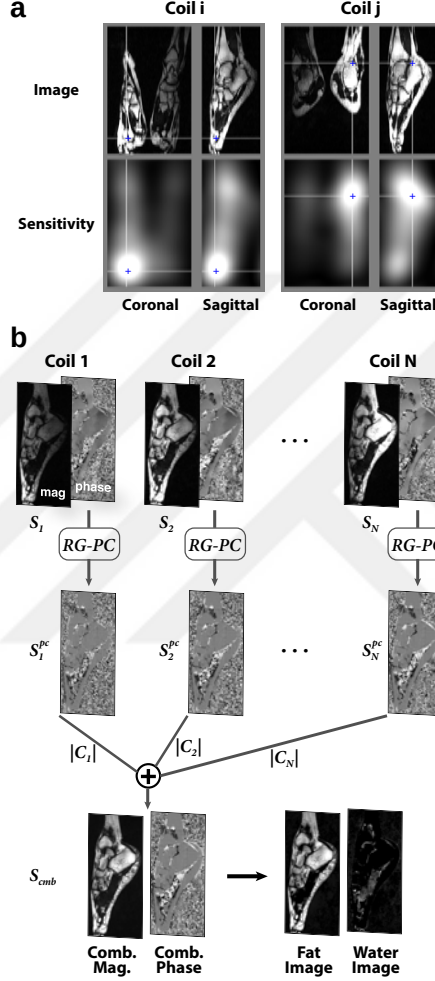


Figure 4.2: **The proposed method.** (a) The initial step of the proposed reconstruction is to identify fat-voxel seeds near the peak of each coil's sensitivity profile. To do this, sensitivity maps are first estimated from the central portion of  $k$ -space data for each coil (bottom row). These maps are then multiplied with the corresponding bSSFP images (top row), and for each coil the voxel that maximizes the product is selected as the starting seed (marked with blue crosses). (b) Following coilwise seed selection, region-growing phase correction (RG-PC) is applied to individual-coil images ( $S_i$ ) in order to remove slowly varying phase components. The phase-corrected coil images ( $S_i^{pc}$ ) are then linearly combined, where images are weighted by the magnitude of respective coil sensitivities ( $|C_i|$ ). Phase-sensitive fat-water separation is finally performed based on the magnitude and phase of the combined image ( $S_{cmb}$ ).

#### 4.2.4 Experiments

Noncontrast-enhanced bSSFP angiograms were acquired at three separate stations in the lower extremities: the thigh, the calf, and the foot. Data were collected on a 1.5T GE (Milwaukee, WI) Signa scanner (40 mT/m maximum strength, 150/T/m.s maximum slew rate) using an 8-channel receive-only phased-array coil. Three volunteers were recruited for the study (two females ages 28 and 34, one male age 28), and all subjects gave written informed consent. The experimental protocols were approved by the local Institutional Review Board in Bilkent University.

To generate angiographic contrast, we used a magnetization-prepared, segmented 3DFT bSSFP sequence that was developed in a recent study [126]. Dual-acquisition bSSFP data were acquired and complex summed to mitigate banding artifacts due to field inhomogeneity [121]. The following parameters were prescribed: superior-inferior readout direction, 90° flip angle, 5 msec TR (the minimum TR allowed by the readout requirements), 2.5 msec TE, 1.7 seconds inversion recovery time, 80 msec  $T_2$ -preparation time, 1200 encodes/segment, 12-tip start-up catalyzation based on a Kaiser-Bessel windowed ramp, and 3 seconds recovery time. A relatively high flip angle was prescribed to increase suppression of muscle signal due to on-resonant magnetization transfer effects. In the thigh, an FOV of 350×350×180 mm<sup>3</sup> was covered with 1.4 mm isotropic resolution in 7 minutes 25 seconds. In the calf, an FOV of 310×240×140 mm<sup>3</sup> was covered with 1 mm isotropic resolution in 7 minutes 46 seconds. In the foot, an FOV of 270×200×140 mm<sup>3</sup> was covered with 1 mm isotropic resolution in a total scan time of 6 minutes 28 seconds. All PS-SSFP reconstructions were performed using a block of size 6×6×6 voxels. Prior to maximum-intensity projections (MIPs), all datasets were transformed to  $k$ -space, zero-padded to double the  $k$ -space coverage, and inverse Fourier-transformed. This procedure was performed to minimize partial volume effects and to improve visualization quality [127].

As a reference for PS-SSFP reconstructions, IDEAL (Iterative Decomposition of Water and Fat with Echo Asymmetry and Least-Squares Estimation) fat–water

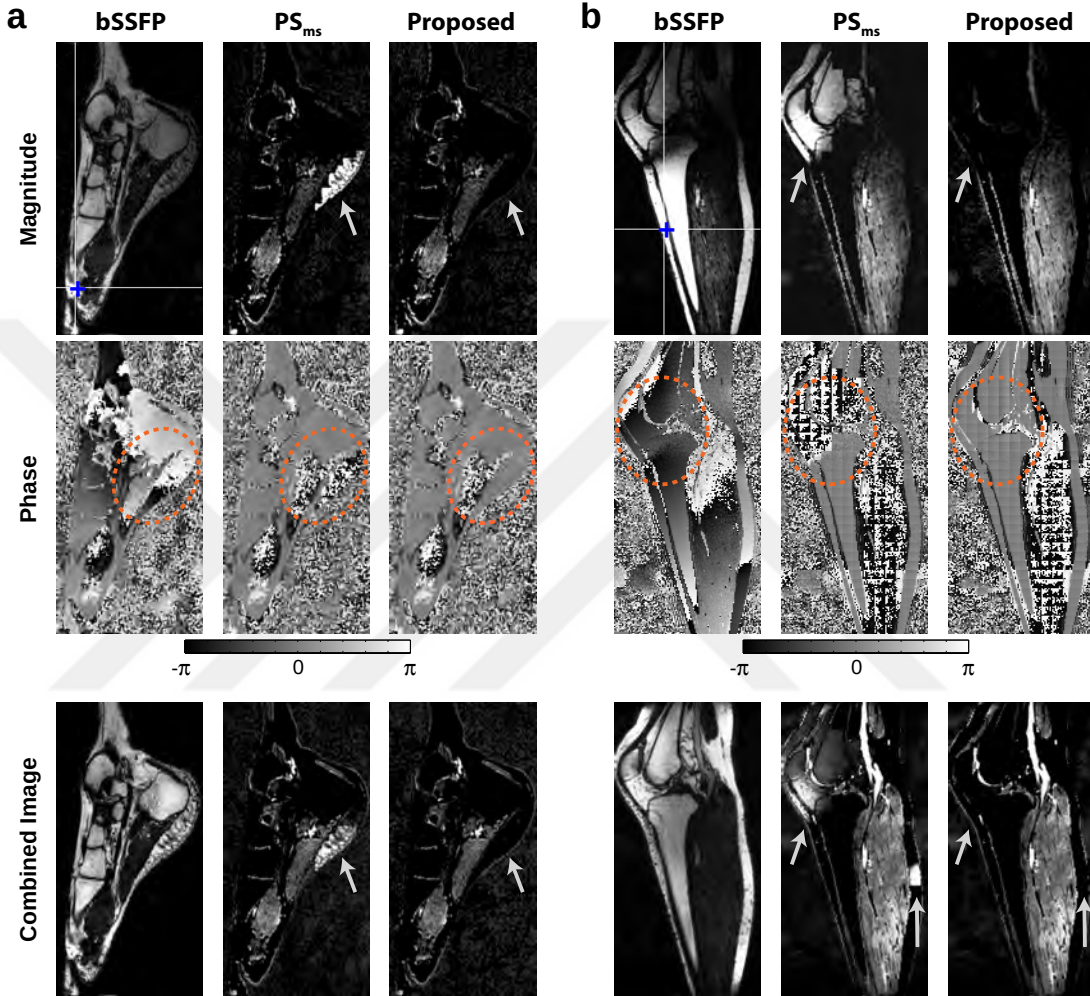


Figure 4.3: **Dual-acquisition bSSFP images in the lower extremities.** The foot (a) and the lower leg (b), collected using an 8-channel phased-array coil. Unseparated complex-sum images are shown for bSSFP, whereas water images are shown for  $PS_{ms}$  (multiseed, individual coil) and the proposed method. Note that  $PS_{ms}$  was performed individually on each coil, and seed selection was identical across the two methods (starting seeds marked with blue crosses). In both a and b, the top row shows magnitude images and the middle row shows phase images (see colorbar) obtained from an individual coil. Meanwhile, the bottom row shows the final image combined across coils. Relatively large variations are observed in phase images away from the sensitivity peak of the coils: in posterior regions of the foot image and superior–anterior regions of the lower leg (marked with dashed ellipses). While these variations cause fat–water misclassification in individual-coil and combined  $PS_{ms}$  images (marked with arrows), the proposed method produces more reliable phase estimates to improve tissue classification.

separation was implemented on dual-acquisition bSSFP data separately collected using a multiecho bipolar-readout 3DFT sequence [128]. The IDEAL sequence was implemented using  $TR = 10$  msec and three echoes with 2.8-msec spacing,  $TE = (1.3, 4.1, 6.9)$  msec. To maintain identical spatial resolution, FOV, and scan time to PS-SSFP acquisitions at each station, IDEAL acquisitions were 1.5-fold accelerated and undersampled data were reconstructed using ARC (Autocalibrating Reconstruction for Cartesian Sampling) [129]. A multipoint IDEAL reconstruction was then performed to separate fat and water signals [130].

Two complementary analyses were performed to evaluate PS-SSFP reconstructions. First, PS-SSFP water images were compared with unseparated bSSFP images. PS-SSFP classifies each voxel as either water or fat. It has been reported in the lower extremities at 1.5T that the average intensity of fat signals is at least twice as high as the average intensity of water signals [125]. Thus, the intensity of PS-SSFP water images should become smaller with improved fat suppression. To compare the level of fat suppression across different PS-SSFP methods, we measured the mean image intensity across axial cross-sections of each reconstruction. In each cross-section we calculated the ratio of the mean intensities measured in PS-SSFP versus bSSFP images ( $R_{bSSFP}$ ; expected to decrease with improved fat suppression).

It is possible that misclassification of water tissue that yield higher bSSFP signals than fat (eg, synovial fluid) can introduce a downward bias in the intensity ratio of PS-SSFP to unseparated bSSFP images. To ensure that these ratio measurements are not significantly biased by suboptimal water signals, a separate control analysis was performed. In this analysis, PS-SSFP and IDEAL water images were compared after intensity normalization. For normalization, an identical region-of-interest (ROI) with homogeneous muscle signal (minimum size of 500 voxels) was selected across all reconstructed volumes. The mean signal intensity within the muscle ROI was normalized to unity separately for PS-SSFP and IDEAL images. PS-SSFP images with more reliable fat–water separation should yield similar intensities to the reference IDEAL images. In each axial cross-section, we calculated the ratio of the mean intensities measured in PS-SSFP

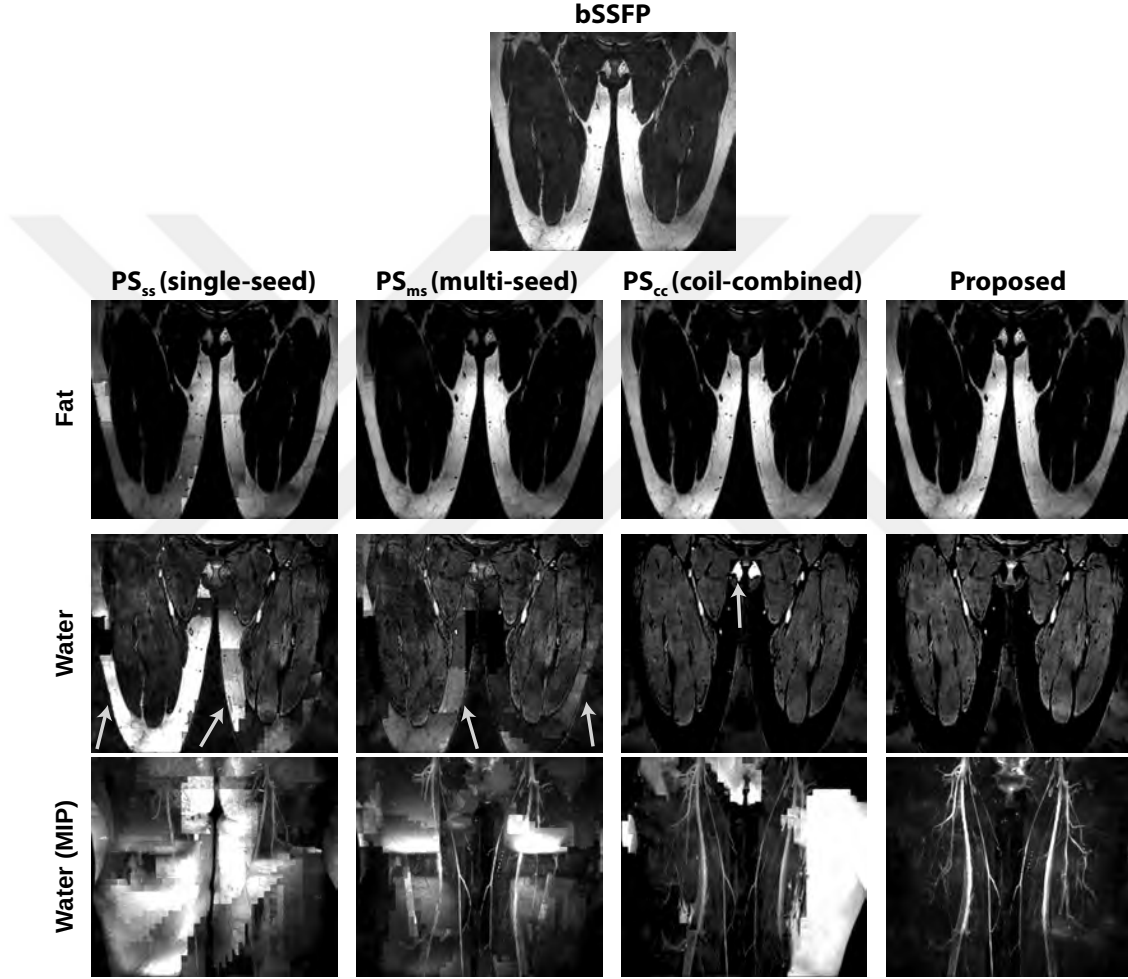


Figure 4.4: **In vivo thigh images combined across 8 channels of a phased-array coil.** First row: The unseparated bSSFP image shown as a reference. Second row: A coronal slice from fat images reconstructed using  $PS_{ss}$ ,  $PS_{ms}$ ,  $PS_{cc}$ , and the proposed method. Third row: A coronal slice from water images reconstructed using  $PS_{ss}$ ,  $PS_{ms}$ ,  $PS_{cc}$ , and the proposed method. Fourth row: Maximum-intensity projections (MIPs) across water images. Residual fat signals are seen in multiple regions of regular PS-SSFP reconstructions (marked with arrows). In contrast, the proposed method maintains reliable fat–water separation across the imaging volume.

versus IDEAL images ( $R_{\text{IDEAL}}$ ; expected to approach 1 with improved water signals). All statistical comparisons were performed using Wilcoxon signed-rank tests ( $P < 0.005$ ).

### 4.3 Results

Balanced SSFP images of the foot and the lower leg from a sample coil, and corresponding water images reconstructed via  $\text{PS}_{\text{ms}}$  (multiseed, individual-coil) and the proposed method, are displayed in Fig. 4.3. The phase images become progressively noisier away from the sensitivity peak for each coil. Increased phase noise in turn causes local failures during phase correction, and regional fat–water misclassification in  $\text{PS}_{\text{ms}}$ . Note that residual fat signals in individual-coil images are prominent even after image combination across coils. In contrast, the proposed method achieves visibly improved fat suppression compared to  $\text{PS}_{\text{ms}}$ .

Representative bSSFP images of the thigh and the lower leg reconstructed via  $\text{PS}_{\text{ss}}$  (single-seed, individual-coil),  $\text{PS}_{\text{ms}}$ , and  $\text{PS}_{\text{cc}}$  (combined-coil) are shown in Figs. 4.4 and 4.5. Both cross-sectional and MIP views of water images—combined across 8 channels of the phased-array coil—show broad regions of fat–water misclassification in  $\text{PS}_{\text{ss}}$ ,  $\text{PS}_{\text{ms}}$ , and  $\text{PS}_{\text{cc}}$ . The combination across coils prior to phase correction in  $\text{PS}_{\text{cc}}$  and the coil-wise seed selection in  $\text{PS}_{\text{ms}}$  partly improve water depiction compared to regular  $\text{PS}_{\text{ss}}$ . Nonetheless, water images obtained via the proposed method demonstrate enhanced tissue separation compared to  $\text{PS}_{\text{cc}}$  and  $\text{PS}_{\text{ms}}$ .

Quantitative assessments of the level of fat suppression ( $R_{\text{bSSFP}}$ ) and the optimality of water signals ( $R_{\text{IDEAL}}$ ) are listed in Table 1. Our proposed method yielded significantly smaller  $R_{\text{bSSFP}}$  than each of the three alternative PS-SSFP reconstructions in the thigh, in the lower leg, and in the foot ( $P < 0.005$ ), indicating that it achieves improved fat suppression. Furthermore, the proposed method attains the most similar image intensities to IDEAL (Fig. 4.6) among

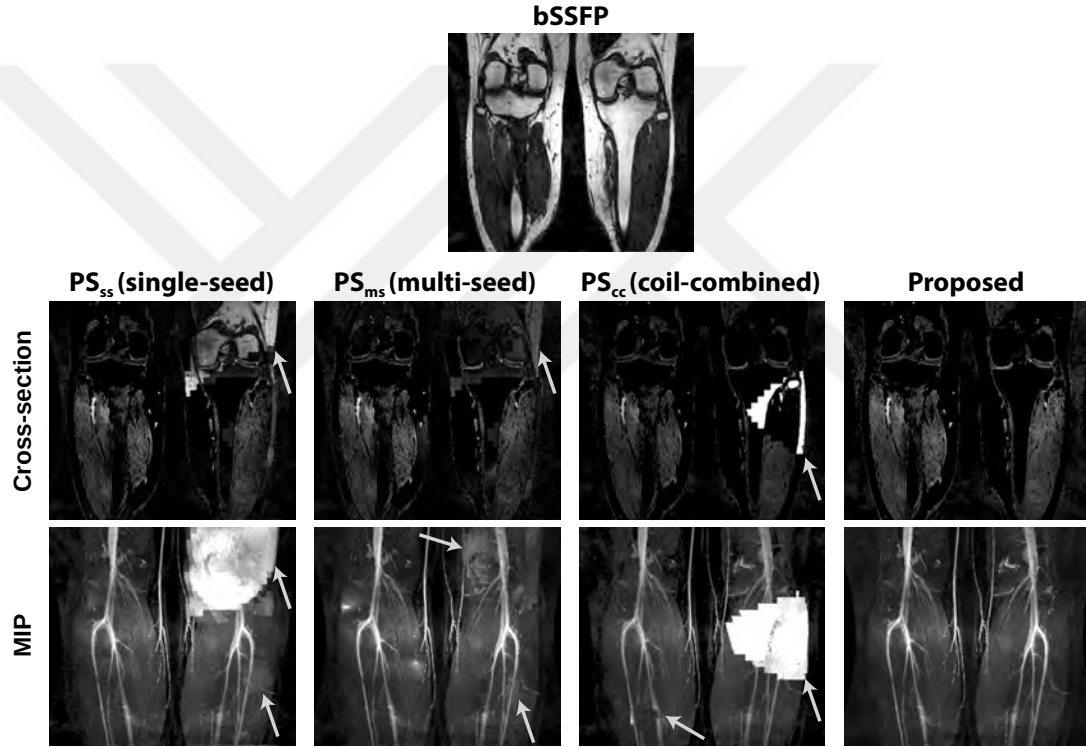


Figure 4.5: **In vivo lower leg images combined across 8 channels of a phased-array coil.** First row: The unseparated bSSFP image shown as a reference. Second row: A coronal slice from water images reconstructed using PS<sub>ss</sub>, PS<sub>ms</sub>, PS<sub>cc</sub>, and the proposed method. Third row: MIPs across water images. Fat–water misclassification is seen broadly across regular PS-SSFP reconstructions (marked with arrows). Meanwhile, the proposed method achieves improved fat–water separation across the entire imaging volume.

all PS-SSFP reconstructions in all body regions ( $P < 0.005$ ). These results indicate that the proposed method achieves significantly more reliable fat–water separation compared to regular PS-SSFP reconstructions.

Table 4.1: **Quantitative Assessments of Fat-Water Separation**

|                    |           | PS <sub>ss</sub> | PS <sub>ms</sub> | PS <sub>cc</sub> | Proposed                 |
|--------------------|-----------|------------------|------------------|------------------|--------------------------|
| R <sub>bSSFP</sub> | Thigh     | 0.363±0.033      | 0.242±0.044      | 0.252±0.042      | 0.181±0.039 <sup>a</sup> |
|                    | Lower leg | 0.290±0.053      | 0.269±0.072      | 0.267±0.088      | 0.246±0.089 <sup>a</sup> |
|                    | Foot      | 0.176±0.046      | 0.174±0.049      | 0.176±0.051      | 0.167±0.048 <sup>a</sup> |
| R <sub>IDEAL</sub> | Thigh     | 2.163±0.555      | 1.346±0.184      | 1.485±0.449      | 1.052±0.213 <sup>a</sup> |
|                    | Lower leg | 1.766±0.493      | 1.586±0.377      | 1.623±0.747      | 1.416±0.403 <sup>a</sup> |
|                    | Foot      | 1.685±0.724      | 1.725±0.782      | 1.729±0.748      | 1.67±0.796 <sup>a</sup>  |

Measurements are reported as mean±SD across cross-sections.

<sup>a</sup>Significantly different results ( $P < 0.005$ ).

## 4.4 Discussion

PS-SSFP fat–water separation employs a correction algorithm to remove phase variations due to undesirable factors including coil sensitivity. This correction proves challenging in phased-array acquisitions since individual-coil image phase typically shows substantial variations away from the coil centers. The proposed method first uses a coil-wise seed selection for individual-coil phase correction and obtains accurate phase estimates near coil-sensitivity peaks. Phase-corrected images are then linearly combined to improve phase estimates in the remaining regions, thereby obtaining enhanced phase estimates across the imaging volume.

A number of effective techniques have been previously proposed for tissue separation in bSSFP imaging [131]. Sequences that employ saturation or spectral-spatial pulses for fat suppression can reduce sensitivity to partial volume effects and chemical-shift artifacts. Yet because these pulses reduce scan efficiency and increase sensitivity to field inhomogeneity, they are often not preferred in imaging extremities [131]. Meanwhile, Dixon-type methods including IDEAL use multi-echo measurements to offer accurate quantification of subvoxel fat–water composition and improved immunity against field inhomogeneity. However, these

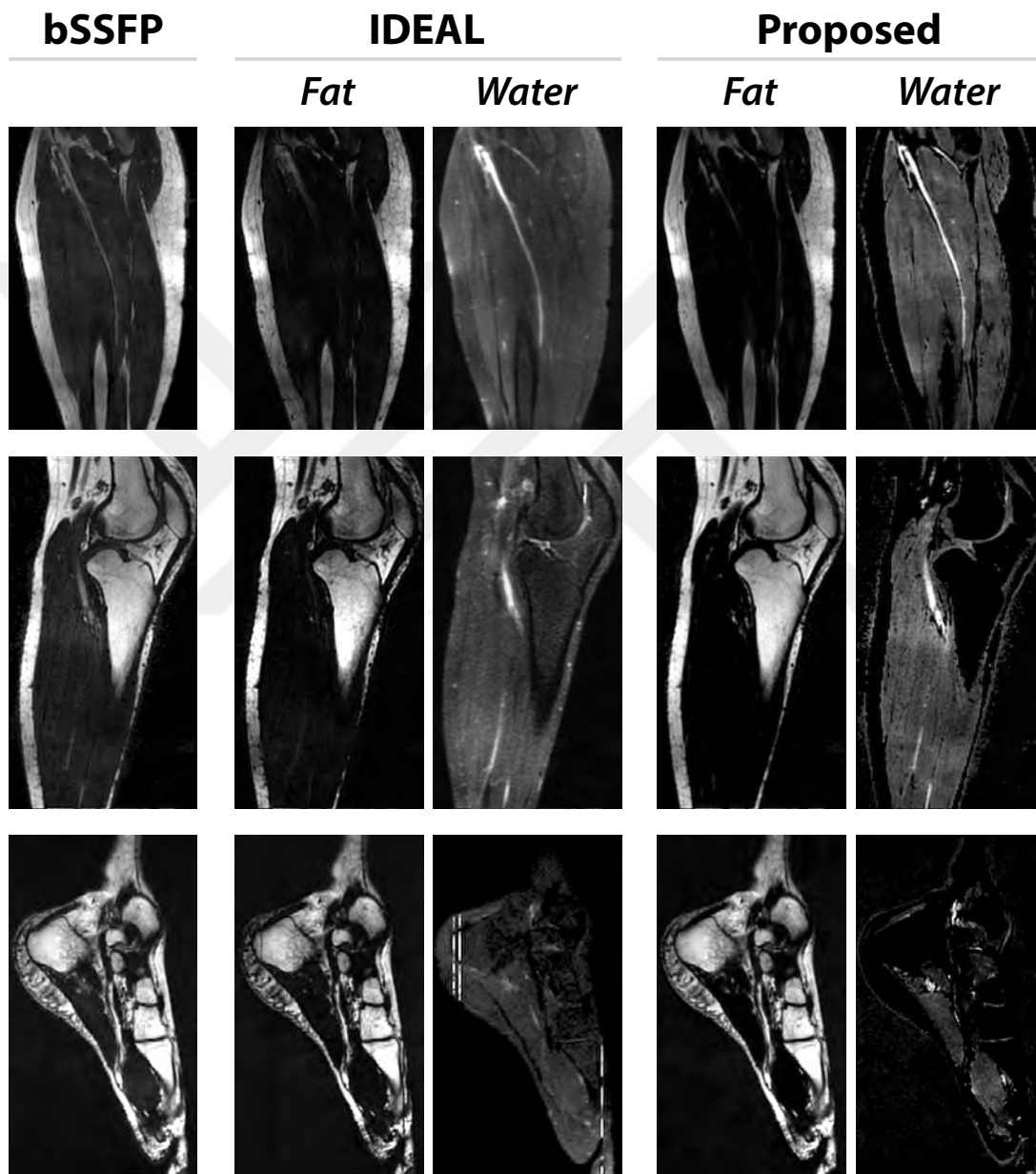


Figure 4.6: Unseparated bSSFP images, fat/water images reconstructed using IDEAL, and fat/water images reconstructed using the proposed method. First row: A sagittal slice from in vivo thigh images. Second row: A sagittal slice from in vivo lower leg images. Third row: A sagittal slice from in vivo foot images.

methods require significantly prolonged scan times and complex reconstruction procedures. In contrast, PS-SSFP does not use special pulses, multiple echoes, or other sequence modifications. It separates fat and water based on a relatively simple phase-sensitive reconstruction, at the expense of increased sensitivity to partial volume effects. Therefore, the proposed method can offer fat–water separation with high scan and processing efficiency.

The application of PS-SSFP on phased-array acquisitions has been considered in several recent studies [22, 122]. One approach is to first form a linear-combination image across coils, and then to perform fat–water separation on the combination [122]. This approach, related to  $PS_{cc}$ , can suffer from residual phase errors in the combined image due to inconsistencies of uncorrected image phase across individual coils. Alternatively, fat–water separation can be performed on individual-coil images as in  $PS_{ss}$  and  $PS_{ms}$  [22]. Images are then combined after correcting for global phase offsets among separate coils [123]. However complex, inter-coil phase inconsistencies may exist even after coil-wise phase correction. The proposed method comprises several technical advances to achieve improved fat–water separation compared to previous approaches. Unlike  $PS_{cc}$ , our method performs phase correction prior to image combination across coils. In this process, it utilizes an automated selection to place seeds in regions of high coil sensitivity as opposed to a common seed for all coils. Finally, unlike  $PS_{ss}$  or  $PS_{ms}$ , it separates fat–water voxels based on an improved phase estimate obtained via combination of corrected coil images.

A basic limitation of PS-SSFP concerns the selection of the sequence TR. A TR of 4.6 msec at 1.5T (or 2.3 msec at 3T) ideally places fat and water resonances at the centers of adjacent passbands. While out-of-phase images can be acquired at odd multiples of the ideal TR (eg, TR = 6.9 msec at 3T), these longer TRs may be undesirable since they increase field-inhomogeneity induced phase variations. However, TR = 6.9 msec or longer could be required for 3T imaging where the ideal TR is too restricted to maintain practical readout resolutions. Therefore, PS-SSFP at higher field strengths can offer improved signal-to-noise ratio (SNR) while increasing sensitivity to field inhomogeneity. It is also possible to prescribe moderately longer/shorter TRs than the ideal value (eg, TR = 5

msec at 1.5T was the minimum possible TR in this study). Such alterations cause the fat resonance to be offset from the passband center, slightly perturbing the fat–water phase difference in dual-acquisition bSSFP. Because PS-SSFP assumes a  $p$ -radians fat–water phase difference, accurate tissue classification is expected when the total phase accrual due to chemical shift and field inhomogeneity is less than  $\pi/2$ -radians. Thus, TR perturbations that increase chemical-shift induced phase effectively reduce the tolerable range of field inhomogeneity.

The proposed method has some other technical limitations that are common to PS-SSFP reconstructions. First, the current implementation comprises two sequential phase-cycled acquisitions, increasing susceptibility to motion. To mitigate artifacts due to patient motion, retrospective motion correction can be performed using navigators incorporated into bSSFP sequences. Second, PS-SSFP classifies each voxel as either fat or water, introducing sensitivity to partial volume effects. Spatial resolution can be increased to reduce this sensitivity, while undersampling can be used to maintain scan-time efficiency. Lastly, when field inhomogeneity is significant, residual banding artifacts might be visible in dual-acquisition bSSFP images. To improve robustness against field inhomogeneity, the proposed method can be applied to complex-sum images from a greater number of phase-cycled acquisitions without any modification.

Several technical developments can be further considered to improve the proposed fat–water separation. While PS-SSFP assumes a single-peak model for the fat resonance, a multipeak extension could offer enhanced tissue delineation [130]. Note that a multipeak model based on single-echo acquisitions would be underconstrained. However, a dual-echo acquisition may enable subvoxel tissue decomposition, assuming that the parameters of the fat spectrum are calibrated a priori [132]. Another improvement concerns the phase correction step of the proposed method. We preferred a region-growing algorithm [22, 120, 121] in this study because it yielded high-quality reconstructions in the lower extremities, and it offered computationally efficient reconstructions for three-dimensional datasets. However, iterative phase-correction algorithms based on graph cuts can further improve reliability against large field inhomogeneities [132].

The proposed fat–water separation method was demonstrated successfully for noncontrast-enhanced angiograms acquired in three body parts in the lower extremities. Although a relatively small number of healthy subjects is reported, these promising results motivate a more direct examination of the reliability of our method under clinical settings. In particular, future studies are warranted that consider a larger population, including patients with vascular disease, various other body parts and habitus, reproducibility assessments, and validation.

In conclusion, we have demonstrated improved phase-sensitive reconstructions of phased-array images, by combining an unsupervised seed-selection procedure for region-growing phase correction and an optimal pooling of phase information across coils. Thus, the proposed method is a promising technique for rapid fat–water separation in bSSFP applications.

## Chapter 5

### Conclusion

In chapter 2, we find that natural semantic search for a word category causes voxelwise tuning shifts toward the target category that cannot be explained by response baseline or gain changes. Tuning shifts occur even when words belonging to target categories are not present in the story. We also find that semantic search task expands the representation of target words while compressing the representation of behaviorally irrelevant words. These results support the idea that attention invokes a matched-filter that dynamically warps representations depending on behavioral goals. This mechanism likely improves task performance during auditory semantic search by more effectively dedicating limited neural resources to the target word categories. Overall, these findings help understand the human ability to perform complex auditory tasks in an ever-changing natural environment.

In chapter 3, we introduced a novel voxelwise modeling approach that leverages stimulus correlations in a natural movie experiment. Results suggest clear improvements in prediction performance, together with increased feature coherence in model weights of cortical voxels. These improvements are best attributed to prevention of crude, unnecessary penalization of model weights, and thereby increased sensitivity in measuring category selectivity of single voxels. Overall, the proposed method can offer improved sensitivity in modeling of single voxels

in naturalistic fMRI experiments.

In the last chapter, we proposed an improved phase estimation method to use in bSSFP applications with multiple coils. The proposed method does not use a modified sequence thus it does not prolong scan times. PS-SSFP fat–water separation employs a correction algorithm to remove phase variations due to undesirable factors including coil sensitivity. This correction proves challenging in phased-array acquisitions since individual-coil image phase typically shows substantial variations away from the coil centers. The proposed method first uses a coil-wise seed selection for individual-coil phase correction and obtains accurate phase estimates near coil-sensitivity peaks. Phase-corrected images are then linearly combined to improve phase estimates in the remaining regions, thereby obtaining enhanced phase estimates across the imaging volume. We demonstrated improved phase-sensitive reconstructions of phased-array images, by combining an unsupervised seed-selection procedure for region-growing phase correction and an optimal pooling of phase information across coils. Thus, the proposed method is a promising technique for rapid fat–water separation in bSSFP applications.

# Bibliography

- [1] A. G. Huth, S. Nishimoto, A. T. Vu, and J. L. Gallant, “A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain,” *Neuron*, vol. 76, no. 6, pp. 1210–1224, 2012.
- [2] S. A. Huettel, A. W. Song, G. McCarthy, *et al.*, *Functional magnetic resonance imaging*, vol. 1. Sinauer Associates Sunderland, MA, 2004.
- [3] K. Grill-Spector and R. Malach, “The human visual cortex,” *Annu. Rev. Neurosci.*, vol. 27, pp. 649–677, 2004.
- [4] B. A. Wandell, S. O. Dumoulin, and A. A. Brewer, “Visual field maps in human cortex,” *Neuron*, vol. 56, no. 2, pp. 366–383, 2007.
- [5] M. Saenz and D. R. M. Langers, “Tonotopic mapping of human auditory cortex,” *Hearing Research*, vol. 307, pp. 42–52, jan 2014.
- [6] C. Humphries, E. Liebenthal, and J. R. Binder, “Tonotopic organization of human auditory cortex,” *NeuroImage*, vol. 50, pp. 1202–11, apr 2010.
- [7] J. M. Thomas, E. Huber, G. C. Stecker, G. M. Boynton, M. Saenz, and I. Fine, “Population receptive field estimates of human auditory cortex,” *NeuroImage*, vol. 105, pp. 428–39, jan 2015.
- [8] K. Grill-Spector, “The neural basis of object perception,” *Current opinion in neurobiology*, vol. 13, no. 2, pp. 159–166, 2003.

- [9] N. Kanwisher, “Functional specificity in the human brain: a window into the functional architecture of the mind,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 25, pp. 11163–11170, 2010.
- [10] P. E. Downing, A.-Y. Chan, M. Peelen, C. Dodds, and N. Kanwisher, “Domain specificity in visual cortex,” *Cerebral cortex*, vol. 16, no. 10, pp. 1453–1461, 2005.
- [11] G. Avidan, U. Hasson, R. Malach, and M. Behrmann, “Detailed exploration of face-related processing in congenital prosopagnosia: 2. functional neuroimaging findings,” *Journal of cognitive neuroscience*, vol. 17, no. 7, pp. 1150–1167, 2005.
- [12] R. F. Schwarzlose, C. I. Baker, and N. Kanwisher, “Separate face and body selectivity on the fusiform gyrus,” *Journal of Neuroscience*, vol. 25, no. 47, pp. 11055–11059, 2005.
- [13] G. McCarthy, A. Puce, J. C. Gore, and T. Allison, “Face-specific processing in the human fusiform gyrus,” *Journal of cognitive neuroscience*, vol. 9, no. 5, pp. 605–610, 1997.
- [14] M. A. Just, V. L. Cherkassky, S. Aryal, and T. M. Mitchell, “A neurosemantic theory of concrete noun representation based on the underlying brain codes,” *PloS one*, vol. 5, no. 1, p. e8622, 2010.
- [15] C. Mummery, K. Patterson, J. Hodges, and C. Price, “Functional neuroanatomy of the semantic system: divisible by what?,” *Journal of cognitive neuroscience*, vol. 10, no. 6, pp. 766–777, 1998.
- [16] A. Caramazza and J. R. Shelton, “Domain-specific knowledge systems in the brain: The animate-inanimate distinction,” *Journal of cognitive neuroscience*, vol. 10, no. 1, pp. 1–34, 1998.
- [17] T. Naselaris, K. N. Kay, S. Nishimoto, and J. L. Gallant, “Encoding and decoding in fmri,” *Neuroimage*, vol. 56, no. 2, pp. 400–410, 2011.
- [18] T. M. Mitchell, S. V. Shinkareva, A. Carlson, K.-M. Chang, V. L. Malave, R. A. Mason, and M. A. Just, “Predicting human brain activity associated

- with the meanings of nouns,” *science*, vol. 320, no. 5880, pp. 1191–1195, 2008.
- [19] D. E. Stansbury, T. Naselaris, and J. L. Gallant, “Natural scene statistics account for the representation of scene categories in human visual cortex,” *Neuron*, vol. 79, no. 5, pp. 1025–1034, 2013.
  - [20] A. G. Huth, W. A. de Heer, T. L. Griffiths, F. E. Theunissen, and J. L. Gallant, “Natural speech reveals the semantic maps that tile human cerebral cortex,” *Nature*, vol. 532, pp. 453–458, apr 2016.
  - [21] O. Bieri, T. C. Mamisch, S. Trattnig, O. Kraff, M. E. Ladd, and K. Scheffler, “Optimized spectrally selective steady-state free precession sequences for cartilage imaging at ultra-high fields,” *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 21, pp. 87–94, 2007.
  - [22] J. P. Wansapura, “Abdominal fat–water separation with SSFP at 3 Tesla,” *Pediatric Radiology*, vol. 37, pp. 69–73, jan 2007.
  - [23] T. Cukur and D. G. Nishimura, “Multiple repetition time balanced steady-state free precession imaging,” *Magnetic resonance in medicine*, vol. 62, pp. 193–204, jul 2009.
  - [24] A. Lu, T. M. Grist, and W. F. Block, “Fat/water separation in single acquisition steady-state free precession using multiple echo radial trajectories,” *Magnetic Resonance in Medicine*, vol. 54, pp. 1051–1057, nov 2005.
  - [25] O. Yilmaz, E. U. Saritas, and T. Çukur, “Enhanced phase-sensitive ssfp reconstruction for fat-water separation in phased-array acquisitions,” *Journal of Magnetic Resonance Imaging*, vol. 44, no. 1, pp. 148–157, 2016.
  - [26] A. R. Mayer, D. Harrington, J. C. Adair, and R. Lee, “The neural networks underlying endogenous auditory covert orienting and reorienting,” *NeuroImage*, vol. 30, pp. 938–949, apr 2006.
  - [27] S. Norman-Haignere, N. Kanwisher, and J. McDermott, “Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition,” *Neuron*, vol. 88, pp. 1281–1296, dec 2015.

- [28] M. Tervaniemi and K. Hugdahl, “Lateralization of auditory-cortex functions,” *Brain Research Reviews*, vol. 43, pp. 231–246, dec 2003.
- [29] A. M. Leaver and J. P. Rauschecker, “Cortical representation of natural complex sounds: effects of acoustic features and auditory object category,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 30, pp. 7604–12, jun 2010.
- [30] R. D. Patterson, S. Uppenkamp, I. S. Johnsrude, and T. D. Griffiths, “The processing of temporal pitch and melody information in auditory cortex,” *Neuron*, vol. 36, pp. 767–76, nov 2002.
- [31] M. Moerel, F. De Martino, E. Formisano, R. Frackowiak, S. Clarke, and M. Saenz, “Processing of natural sounds in human auditory cortex: tonotopy, spectral tuning, and relation to voice sensitivity,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 32, pp. 14205–16, oct 2012.
- [32] J. R. Kerlin, A. J. Shahin, and L. M. Miller, “Attentional Gain Control of Ongoing Cortical Speech Representations in a ”Cocktail Party”,” *Journal of Neuroscience*, vol. 30, pp. 620–628, jan 2010.
- [33] B. Morillon, C. E. Schroeder, and V. Wyart, “Motor contributions to the temporal precision of auditory attention,” *Nature Communications*, vol. 5, p. 5255, oct 2014.
- [34] T. M. Teshiba, J. Ling, D. A. Ruhl, B. S. Bedrick, A. Peña, and A. R. Mayer, “Evoked and Intrinsic Asymmetries during Auditory Attention: Implications for the Contralateral and Neglect Models of Functioning,” *Cerebral Cortex*, vol. 23, pp. 560–569, mar 2013.
- [35] A. Seydell-Greenwald, A. S. Greenberg, and J. P. Rauschecker, “Are you listening? Brain activation associated with sustained nonspatial auditory attention in the presence and absence of stimulation,” *Human Brain Mapping*, vol. 35, no. 5, pp. 2233–2252, 2014.

- [36] S. Shomstein and S. Yantis, “Parietal Cortex Mediates Voluntary Control of Spatial and Nonspatial Auditory Attention,” *Journal of Neuroscience*, vol. 26, pp. 435–439, jan 2006.
- [37] L. Kong, S. W. Michalka, M. L. Rosen, S. L. Sheremata, J. D. Swisher, B. G. Shinn-Cunningham, and D. C. Somers, “Auditory spatial attention representations in the human cerebral cortex,” *Cerebral Cortex*, vol. 24, no. 3, pp. 773–784, 2014.
- [38] K. V. Nourski, M. Steinschneider, A. E. Rhone, and M. A. Howard III, “Intracranial Electrophysiology of Auditory Selective Attention Associated with Speech Classification Tasks,” *Frontiers in Human Neuroscience*, vol. 10, no. January, p. 691, 2017.
- [39] M. Steinschneider, K. V. Nourski, A. E. Rhone, H. Kawasaki, H. Oya, and M. A. Howard, “Differential activation of human core, non-core and auditory-related cortex during speech categorization tasks as revealed by intracranial recordings,” *Frontiers in Neuroscience*, vol. 8, p. 240, aug 2014.
- [40] A. J. Shahin, C. Alain, and T. W. Picton, “Scalp topography and intracerebral sources for ERPs recorded during auditory target detection,” *Brain Topography*, vol. 19, pp. 89–105, dec 2006.
- [41] S. V. David, B. Y. Hayden, J. A. Mazer, and J. L. Gallant, “Attention to Stimulus Features Shifts Spectral Tuning of V4 Neurons during Natural Vision,” *Neuron*, vol. 59, pp. 509–521, aug 2008.
- [42] C. J. McAdams and J. H. Maunsell, “Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 19, pp. 431–41, jan 1999.
- [43] A. Compte and X.-J. Wang, “Tuning Curve Shift by Attention Modulation in Cortical Neurons: a Computational Study of its Mechanisms,” *Cerebral Cortex*, vol. 16, pp. 761–778, jun 2006.

- [44] B. A. Olshausen, C. H. Anderson, and D. C. Van Essen, “A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information.,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 13, pp. 4700–19, nov 1993.
- [45] J. Fritz, S. Shamma, M. Elhilali, and D. Klein, “Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex,” *Nature Neuroscience*, vol. 6, pp. 1216–1223, nov 2003.
- [46] P. Lakatos, A. Barczak, S. A. Neymotin, T. McGinnis, D. Ross, D. C. Javitt, and M. N. O’Connell, “Global dynamics of selective attention and its lapses in primary auditory cortex,” *Nature Neuroscience*, vol. 19, pp. 1707–1717, sep 2016.
- [47] P. Lakatos, G. Musacchia, M. N. O’Connel, A. Y. Falchier, D. C. Javitt, and C. E. Schroeder, “The Spectrotemporal Filter Mechanism of Auditory Selective Attention,” *Neuron*, vol. 77, no. 4, pp. 750–761, 2013.
- [48] S. Da Costa, W. van der Zwaag, L. M. Miller, S. Clarke, and M. Saenz, “Tuning In to Sound: Frequency-Selective Attentional Filter in Human Primary Auditory Cortex,” *Journal of Neuroscience*, vol. 33, no. 5, pp. 1858–1863, 2013.
- [49] J. Ahveninen, I. P. Jaaskelainen, T. Raij, G. Bonmassar, S. Devore, M. Hamalainen, S. Levanen, F.-H. H. Lin, M. Sams, B. G. Shinn-Cunningham, T. Witzel, J. W. Belliveau, I. P. Jääskeläinen, M. Hämäläinen, and S. Levänen, “Task-modulated ”what” and ”where” pathways in human auditory cortex,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 39, pp. 14608–14613, 2006.
- [50] C. I. Petkov, X. Kang, K. Alho, O. Bertrand, E. W. Yund, and D. L. Woods, “Attentional modulation of human auditory cortex,” *Nature Neuroscience*, vol. 7, pp. 658–663, jun 2004.

- [51] N. Mesgarani and E. F. Chang, “Selective cortical representation of attended speaker in multi-talker speech perception,” *Nature*, vol. 485, pp. 233–236, apr 2012.
- [52] H. Renvall, N. Staeren, C. S. Barz, A. Ley, and E. Formisano, “Attention Modulates the Auditory Cortical Processing of Spatial and Category Cues in Naturalistic Auditory Scenes,” *Frontiers in Neuroscience*, vol. 10, p. 254, jun 2016.
- [53] T. Çukur, S. Nishimoto, A. G. Huth, and J. L. Gallant, “Attention during natural vision warps semantic representation across the human brain,” *Nature Neuroscience*, vol. 16, pp. 763–770, apr 2013.
- [54] K. N. Kay, T. Naselaris, R. J. Prenger, and J. L. Gallant, “Identifying natural images from human brain activity,” *Nature*, vol. 452, pp. 352–355, mar 2008.
- [55] N. Kanwisher, J. McDermott, and M. M. Chun, “The fusiform face area: a module in human extrastriate cortex specialized for face perception,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 17, pp. 4302–11, jun 1997.
- [56] P. E. Downing, Y. Jiang, M. Shuman, and N. Kanwisher, “A cortical area selective for visual processing of the human body,” *Science*, vol. 293, pp. 2470–2473, sep 2001.
- [57] R. Epstein and N. Kanwisher, “A cortical representation of the local visual environment,” *Nature*, vol. 392, pp. 598–601, apr 1998.
- [58] K. Nakamura, “Functional delineation of the human occipito-temporal areas related to face and scene processing: A PET study,” *Brain*, vol. 123, pp. 1903–1912, sep 2000.
- [59] M. W. Woolrich, S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, T. Behrens, C. Beckmann, M. Jenkinson, and S. M. Smith, “Bayesian analysis of neuroimaging data in FSL,” *NeuroImage*, vol. 45, pp. S173–S186, mar 2009.

- [60] T. D. Verstynen and V. Deshpande, “Using pulse oximetry to account for high and low frequency physiological artifacts in the BOLD signal,” *NeuroImage*, vol. 55, pp. 1633–1644, apr 2011.
- [61] P. D. Welch, “The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms,” *IEEE Transactions on Audio and Electroacoustics*, vol. 15, pp. 70–73, jun 1967.
- [62] A. M. Dale, B. Fischl, and M. I. Sereno, “Cortical surface-based analysis: I. Segmentation and surface reconstruction,” *NeuroImage*, vol. 9, pp. 179–194, feb 1999.
- [63] J. S. Gao, A. G. Huth, M. D. Lescroart, and J. L. Gallant, “Pycortex: an interactive surface visualizer for fMRI,” *Frontiers in Neuroinformatics*, vol. 9, p. 23, sep 2015.
- [64] C. E. Connor, D. C. Preddie, J. L. Gallant, and D. C. Van Essen, “Spatial attention effects in macaque area V4,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 17, pp. 3201–14, may 1997.
- [65] K. Johnston and S. Everling, “Neural Activity in Monkey Prefrontal Cortex Is Modulated by Task Context and Behavioral Instruction during Delayed-match-to-sample and Conditional Prosaccade-Antisaccade Tasks,” *Journal of Cognitive Neuroscience*, vol. 18, pp. 749–765, may 2006.
- [66] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, “The brain’s default network: Anatomy, function, and relevance to disease,” *Annals of the New York Academy of Sciences*, vol. 1124, pp. 1–38, mar 2008.
- [67] P. Vetter, F. W. Smith, and L. Muckli, “Decoding sound and imagery content in early visual cortex,” *Current Biology*, vol. 24, no. 11, pp. 1256–1262, 2014.

- [68] M. Vigneau, V. Beaucousin, P.-Y. Herve, H. Duffau, F. Crivello, O. Houde, B. Mazoyer, and N. Tzourio-Mazoyer, “Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing,” *Neuroimage*, vol. 30, no. 4, pp. 1414–1432, 2006.
- [69] C. J. Price, “The anatomy of language: Contributions from functional neuroimaging,” *Journal of Anatomy*, vol. 197, pp. 335–359, oct 2000.
- [70] C. Grady, J. V. Meter, and J. Maisog, “Attention-related modulation of activity in primary and secondary auditory cortex,” *NeuroReport*, vol. 8, pp. 2511–2516, jul 1997.
- [71] J. V. Haxby, “Multivariate pattern analysis of fmri: the early beginnings,” *Neuroimage*, vol. 62, no. 2, pp. 852–855, 2012.
- [72] S. Haufe, F. Meinecke, K. Görgen, S. Dähne, J.-D. Haynes, B. Blankertz, and F. Bießmann, “On the interpretation of weight vectors of linear models in multivariate neuroimaging,” *Neuroimage*, vol. 87, pp. 96–110, 2014.
- [73] T. Naselaris, R. J. Prenger, K. N. Kay, M. Oliver, and J. L. Gallant, “Bayesian reconstruction of natural images from human brain activity,” *Neuron*, vol. 63, pp. 902–15, sep 2009.
- [74] M. D. Lescroart, D. E. Stansbury, and J. L. Gallant, “Fourier power, subjective distance, and object categories all provide plausible models of BOLD responses in scene-selective visual areas,” *Frontiers in Computational Neuroscience*, vol. 9, p. 135, nov 2015.
- [75] S. Nishimoto, A. T. Vu, T. Naselaris, Y. Benjamini, B. Yu, and J. L. Gallant, “Reconstructing visual experiences from brain activity evoked by natural movies,” *Current biology : CB*, vol. 21, pp. 1641–6, oct 2011.
- [76] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2013.
- [77] E. Erwin, K. Obermayer, and K. Schulten, “Models of orientation and ocular dominance columns in the visual cortex: a critical comparison,” *Neural computation*, vol. 7, pp. 425–68, may 1995.

- [78] E. Zarahn, G. K. Aguirre, and M. D’Esposito, “Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions.,” *NeuroImage*, vol. 5, pp. 179–97, apr 1997.
- [79] A. O. Nunez-Elizalde, A. G. Huth, and J. L. Gallant, “Voxelwise encoding models with non-spherical multivariate normal priors,” *bioRxiv*, p. 386318, aug 2018.
- [80] E. Çelik, S. U. H. Dar, Ö. Yılmaz, Ü. Keleş, and T. Çukur, “Spatially informed voxelwise modeling for naturalistic fMRI experiments,” *NeuroImage*, vol. 186, pp. 741–757, feb 2019.
- [81] A. E. Hoerl and R. W. Kennard, “Ridge Regression: Biased Estimation for Nonorthogonal Problems,” *Technometrics*, vol. 12, pp. 55–67, feb 1970.
- [82] D. C. Van Essen, H. A. Drury, S. Joshi, and M. I. Miller, “Functional and structural mapping of human cerebral cortex: solutions are in the surfaces.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, pp. 788–95, feb 1998.
- [83] A. Tucholka, V. Fritsch, J.-B. Poline, and B. Thirion, “An empirical comparison of surface-based and volume-based group studies in neuroimaging,” *NeuroImage*, vol. 63, pp. 1443–1453, nov 2012.
- [84] T. Çukur, A. G. Huth, S. Nishimoto, and J. L. Gallant, “Functional subdomains within human FFA.,” *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 33, no. 42, pp. 16748–66, 2013.
- [85] G. A. Miller, “Wordnet: A lexical database for english,” *Commun. ACM*, vol. 38, pp. 39–41, Nov. 1995.
- [86] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, and R. Harshman, “Indexing by latent semantic analysis,” *Journal of the American Society for Information Science*, vol. 41, pp. 391–407, sep 1990.
- [87] K. Lund and C. Burgess, “Producing high-dimensional semantic spaces from lexical co-occurrence,” *Behavior Research Methods, Instruments, & Computers*, vol. 28, pp. 203–208, jun 1996.

- [88] T. M. Mitchell, S. V. Shinkareva, A. Carlson, K.-M. Chang, V. L. Malave, R. A. Mason, and M. A. Just, “Predicting Human Brain Activity Associated with the Meanings of Nouns,” *Science*, vol. 320, pp. 1191–1195, may 2008.
- [89] P. D. Turney and P. Pantel, “From frequency to meaning: Vector space models of semantics,” *Journal of Artificial Intelligence Research*, vol. 37, pp. 141–188, mar 2010.
- [90] L. Wehbe, B. Murphy, P. Talukdar, A. Fyshe, A. Ramdas, and T. Mitchell, “Simultaneously Uncovering the Patterns of Brain Regions Involved in Different Story Reading Subprocesses,” *PLoS ONE*, vol. 9, p. e112575, nov 2014.
- [91] S. A. Engel, G. H. Glover, and B. A. Wandell, “Retinotopic organization in human visual cortex and the spatial precision of functional MRI,” *Cerebral cortex (New York, N.Y. : 1991)*, vol. 7, pp. 181–92, mar 1997.
- [92] R. B. Tootell, N. K. Hadjikhani, W. Vanduffel, A. K. Liu, J. D. Mendola, M. I. Sereno, and A. M. Dale, “Functional analysis of primary visual cortex (V1) in humans,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, pp. 811–7, feb 1998.
- [93] S. V. David and J. L. Gallant, “Predicting neuronal responses during natural vision,” *Network: Computation in Neural Systems*, vol. 16, pp. 239–260, jan 2005.
- [94] A. Hsu, A. Borst, and F. E. Theunissen, “Quantifying variability in neural responses and its application for the validation of model predictions,” *Network (Bristol, England)*, vol. 15, pp. 91–109, may 2004.
- [95] M. Sahani and J. F. Linden, “How Linear are Auditory Cortical Responses?,” 2003.
- [96] A. C. Connolly, J. S. Guntupalli, J. Gors, M. Hanke, Y. O. Halchenko, Y.-C. Wu, H. Abdi, and J. V. Haxby, “The Representation of Biological Classes in the Human Brain,” *Journal of Neuroscience*, vol. 32, no. 8, 2012.

- [97] P. E. Downing, A. W.-Y. Chan, M. V. Peelen, C. M. Dodds, and N. Kanwisher, “Domain specificity in visual cortex,” *Cerebral cortex (New York, N.Y. : 1991)*, vol. 16, pp. 1453–61, oct 2006.
- [98] H. P. Op De Beeck, J. Haushofer, and N. G. Kanwisher, “Interpreting fMRI data: Maps, modules and dimensions,” feb 2008.
- [99] T. Tran, D. Phung, W. Luo, and S. Venkatesh, “Stabilized sparse ordinal regression for medical risk stratification,” *Knowledge and Information Systems*, vol. 43, pp. 555–582, jun 2015.
- [100] X. Zhu, H.-I. Suk, L. Wang, S.-W. Lee, D. Shen, and Alzheimer’s Disease Neuroimaging Initiative, “A novel relational regularization feature selection method for joint regression and classification in AD diagnosis,” *Medical Image Analysis*, vol. 38, pp. 205–214, may 2017.
- [101] T. Sandler, J. Blitzer, P. Pratim Talukdar, and L. Ungar, *Regularized Learning with Networks of Features*. jan 2008.
- [102] B. Zhang and M. Ostendorf, “Semi-supervised learning for text classification using feature affinity regularization,” in *Symposium on Machine Learning in Speech and Language Processing*, 2012.
- [103] A. O. Nunez-Elizalde, A. G. Huth, and J. L. Gallant, “Voxelwise encoding models with non-spherical multivariate normal priors,” *NeuroImage*, vol. 197, pp. 482–492, 2019.
- [104] A. Clarke and L. K. Tyler, “Understanding What We See: How We Derive Meaning From Vision,” *Trends in cognitive sciences*, vol. 19, pp. 677–687, nov 2015.
- [105] T. A. Carlson, R. A. Simmons, N. Kriegeskorte, and L. R. Slevc, “The emergence of semantic meaning in the ventral temporal pathway,” *Journal of cognitive neuroscience*, vol. 26, pp. 120–31, jan 2014.
- [106] H. Wen, J. Shi, W. Chen, and Z. Liu, “Deep Residual Network Predicts Cortical Representation and Organization of Visual Features for Rapid Categorization,” *Scientific reports*, vol. 8, no. 1, p. 3752, 2018.

- [107] T. K. Landauer, “Latent Semantic Analysis,” in *Encyclopedia of Cognitive Science*, Chichester: John Wiley & Sons, Ltd, jan 2006.
- [108] J. Pennington, R. Socher, and C. Manning, “Glove: Global Vectors for Word Representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (Stroudsburg, PA, USA), pp. 1532–1543, Association for Computational Linguistics, 2014.
- [109] T. Mikolov, W.-t. Yih, and G. Zweig, “Linguistic Regularities in Continuous Space Word Representations,” 2013.
- [110] T. Çukur, J. H. Lee, N. K. Bangerter, B. A. Hargreaves, and D. G. Nishimura, “Non-contrast-enhanced flow-independent peripheral MR angiography with balanced SSFP,” *Magnetic Resonance in Medicine*, vol. 61, pp. 1533–1539, jun 2009.
- [111] J. Leupold, J. Hennig, and K. Scheffler, “Alternating repetition time balanced steady state free precession,” *Magnetic Resonance in Medicine*, vol. 55, pp. 557–565, mar 2006.
- [112] J. Absil, V. Denolin, and T. Metens, “Fat attenuation using a dual steady-state balanced-SSFP sequence with periodically variable flip angles,” *Magnetic Resonance in Medicine*, vol. 55, pp. 343–351, feb 2006.
- [113] K. Scheffler, O. Heid, and J. Hennig, “Magnetization preparation during the steady state: fat-saturated 3D TrueFISP,” *Magnetic resonance in medicine*, vol. 45, pp. 1075–80, jun 2001.
- [114] D. Paul, J. Hennig, and M. Zaitsev, “Intrinsic fat suppression in TIDE balanced steady-state free precession imaging,” *Magnetic Resonance in Medicine*, vol. 56, pp. 1328–1335, dec 2006.
- [115] J. A. Derbyshire, D. A. Herzka, and E. R. McVeigh, “S5FP: Spectrally selective suppression with steady state free precession,” *Magnetic Resonance in Medicine*, vol. 54, pp. 918–928, oct 2005.
- [116] T.-Y. Huang, H.-W. Chung, F.-N. Wang, C.-W. Ko, and C.-Y. Chen, “Fat and water separation in balanced steady-state free precession using the

- Dixon method,” *Magnetic Resonance in Medicine*, vol. 51, pp. 243–247, feb 2004.
- [117] S. B. Reeder, M. Markl, H. Yu, J. C. Hellinger, R. J. Herfkens, and N. J. Pelc, “Cardiac CINE imaging with IDEAL water-fat separation and steady-state free precession,” *Journal of Magnetic Resonance Imaging*, vol. 22, pp. 44–52, jul 2005.
  - [118] T. Cukur and D. G. Nishimura, “Fat-water separation with alternating repetition time balanced SSFP,” *Magnetic resonance in medicine*, vol. 60, pp. 479–84, aug 2008.
  - [119] J. Ma, “Breath-hold water and fat imaging using a dual-echo two-point dixon technique with an efficient and robust phase-correction algorithm,” *Magnetic Resonance in Medicine*, vol. 52, pp. 415–419, aug 2004.
  - [120] B. A. Hargreaves, S. S. Vasanawala, K. S. Nayak, B. S. Hu, and D. G. Nishimura, “Fat-suppressed steady-state free precession imaging using phase detection,” *Magnetic Resonance in Medicine*, vol. 50, pp. 210–213, jul 2003.
  - [121] B. A. Hargreaves, N. K. Bangerter, A. Shimakawa, S. S. Vasanawala, J. H. Brittain, and D. G. Nishimura, “Dual-acquisition phase-sensitive fat–water separation using balanced steady-state free precession,” *Magnetic Resonance Imaging*, vol. 24, pp. 113–122, feb 2006.
  - [122] S. S. Vasanawala, B. A. Hargreaves, and D. G. Nishimura, “Phase Sensitive SSFP Parallel Imaging,” Tech. Rep. 2, 1999.
  - [123] J. Ma, “Multislice and multicoil phase-sensitive inversion-recovery imaging,” *Magnetic Resonance in Medicine*, vol. 53, pp. 904–910, apr 2005.
  - [124] T. Cukur, M. Lustig, and D. G. Nishimura, “Multiple-profile homogeneous image combination: application to phase-cycled SSFP and multicoil imaging,” *Magnetic resonance in medicine*, vol. 60, pp. 732–8, sep 2008.

- [125] A. Mazumdar, B. Hargreaves, E. Han, A. Brau, H. Yu, and J. Brittain, “Automated fat-water identification in phase sensitive ssfp,” in *Proc 13th Annual Meeting ISMRM, Miami*, p. 2307, 2005.
- [126] T. Çukur, A. Shimakawa, H. Yu, B. A. Hargreaves, B. S. Hu, D. G. Nishimura, and J. H. Brittain, “Magnetization-prepared IDEAL bSSFP: A flow-independent technique for noncontrast-enhanced peripheral angiography,” *Journal of Magnetic Resonance Imaging*, vol. 33, pp. 931–939, apr 2011.
- [127] Y. P. Du, D. L. Parker, W. L. Davis, and G. Cao, “Reduction of partial-volume artifacts with zero-filled interpolation in three-dimensional MR angiography,” *Journal of magnetic resonance imaging : JMRI*, vol. 4, no. 5, pp. 733–41.
- [128] W. Lu, H. Yu, A. Shimakawa, M. Alley, S. B. Reeder, and B. A. Hargreaves, “Water-fat separation with bipolar multiecho sequences,” *Magnetic Resonance in Medicine*, vol. 60, pp. 198–209, jul 2008.
- [129] A. C. Brau, P. J. Beatty, S. Skare, and R. Bammer, “Comparison of reconstruction accuracy and efficiency among autocalibrating data-driven parallel imaging methods,” *Magnetic Resonance in Medicine*, vol. 59, pp. 382–395, feb 2008.
- [130] H. Yu, A. Shimakawa, C. A. McKenzie, E. Brodsky, J. H. Brittain, and S. B. Reeder, “Multiecho water-fat separation and simultaneous r estimation with multifrequency fat spectrum modeling,” *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 60, no. 5, pp. 1122–1134, 2008.
- [131] T. A. Bley, O. Wieben, C. J. François, J. H. Brittain, and S. B. Reeder, “Fat and water magnetic resonance imaging,” *Journal of Magnetic Resonance Imaging*, vol. 31, pp. 4–18, jan 2010.
- [132] D. Hernando, P. Kellman, J. P. Haldar, and Z.-P. Liang, “Robust water/fat separation in the presence of large field inhomogeneities using a graph cut

algorithm,” *Magnetic Resonance in Medicine*, vol. 63, pp. NA–NA, jan  
2009.



# Appendix A

## Supplementary Figures-1

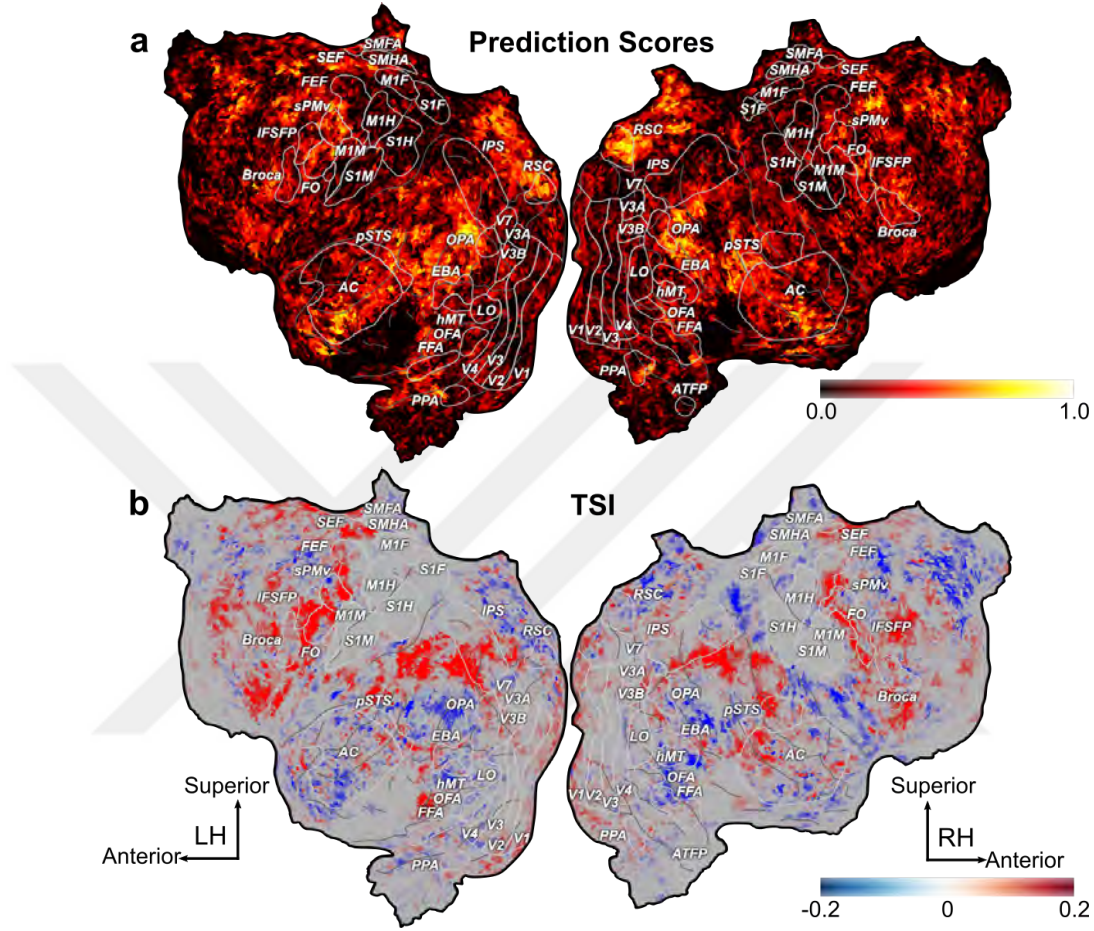


Figure A.1: **Cortical flat maps of prediction score and TSI for subject S1.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S1. Prediction scores were calculated as Pearson's correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) are accurately predicted by the semantic model. (b) Cortical distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing.



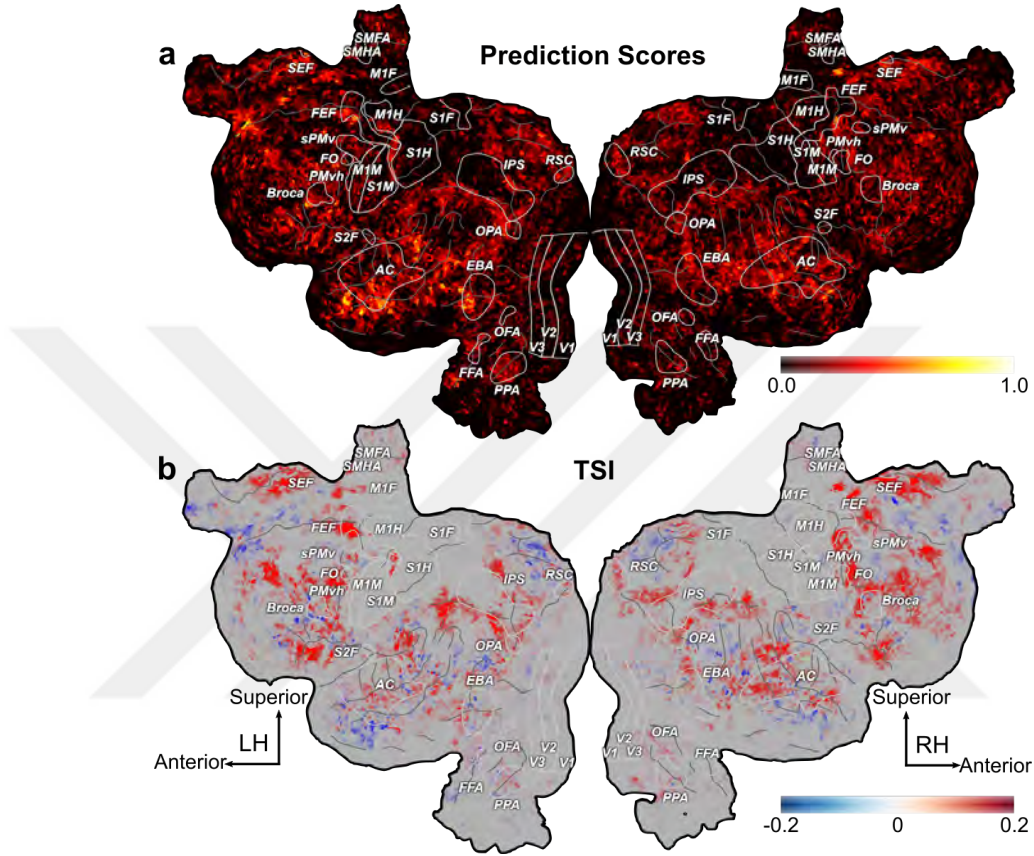


Figure A.3: **Cortical flat maps of prediction score and TSI for subject S3.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S3. Prediction scores were calculated as Pearson's correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) were accurately predicted by the semantic model. (b) distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing.

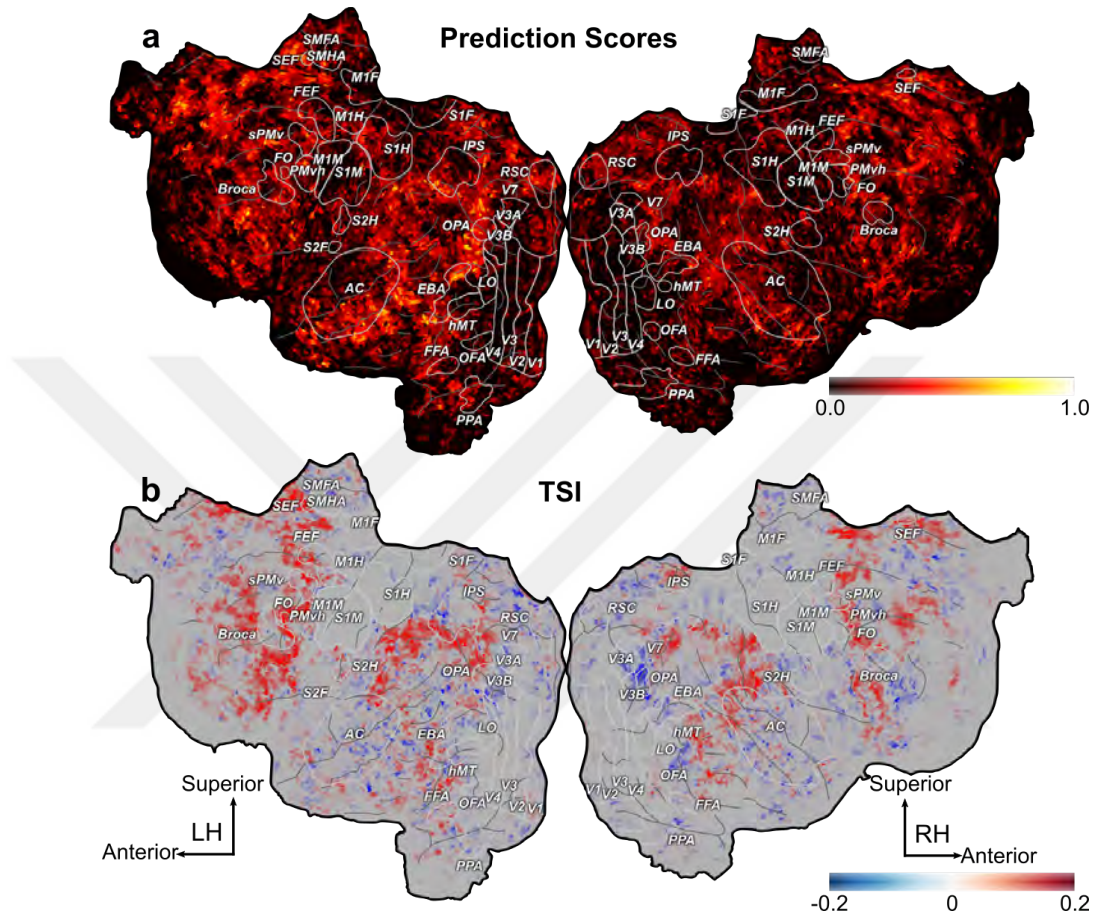


Figure A.4: **Cortical flat maps of prediction score and TSI for subject S4.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S4. Prediction scores were calculated as Pearson's correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) were accurately predicted by the semantic model. (b) Cortical distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing.

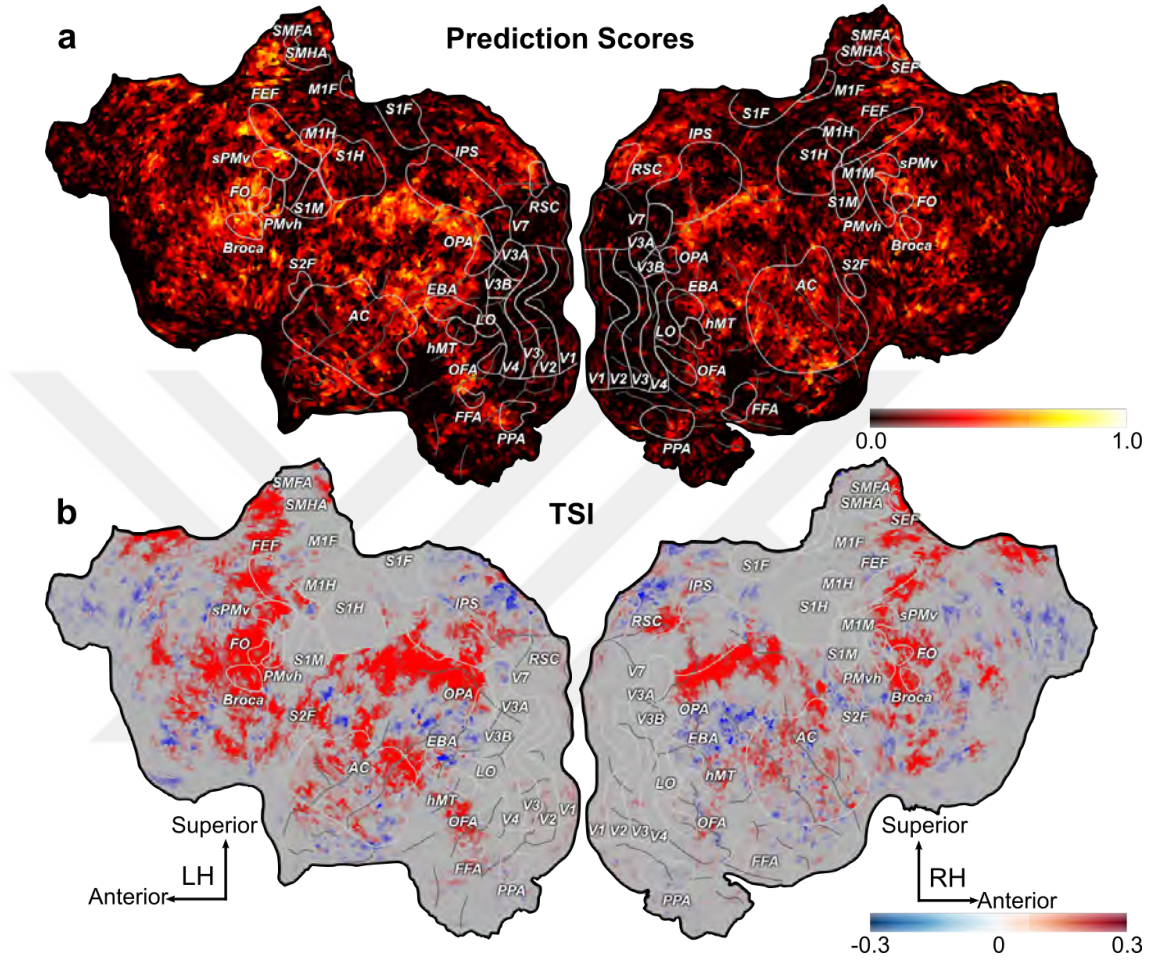


**Figure A.5: Cortical flat maps of prediction score and TSI for subject S5.** (a) To assess the functionality of semantic model, the prediction scores were visualized on cortical flat map for subject S5. Prediction scores were calculated as Pearson's correlation coefficient between the predicted and measured BOLD responses. Regions of fMRI signal dropout are shown with black patches and the boundaries of regions-of-interest (white), gyri (black) and sulci (black) are indicated by lines. Yellow voxels have higher prediction scores compared to black voxels as shown in color bar. Results indicate that many voxels in brain areas, including lateral temporal cortex (LTC), ventral temporal cortex (VTC), lateral parietal cortex (LPC), medial parietal cortex (MPC), medial prefrontal cortex, superior prefrontal cortex (SPFC), and inferior prefrontal cortex (IPFC) were accurately predicted by the semantic model. (b) Cortical distribution of TSIs were visualized on the cortical flat map. Poorly modeled voxels are masked by means of their prediction score, and attentional modulation significance (see Methods), so that well-modeled and significantly attention-modulated voxels become more visible. Results indicate that voxels of similar TSI values are distributed as small clusters or stripes mainly across regions that take place in attention control and language processing.

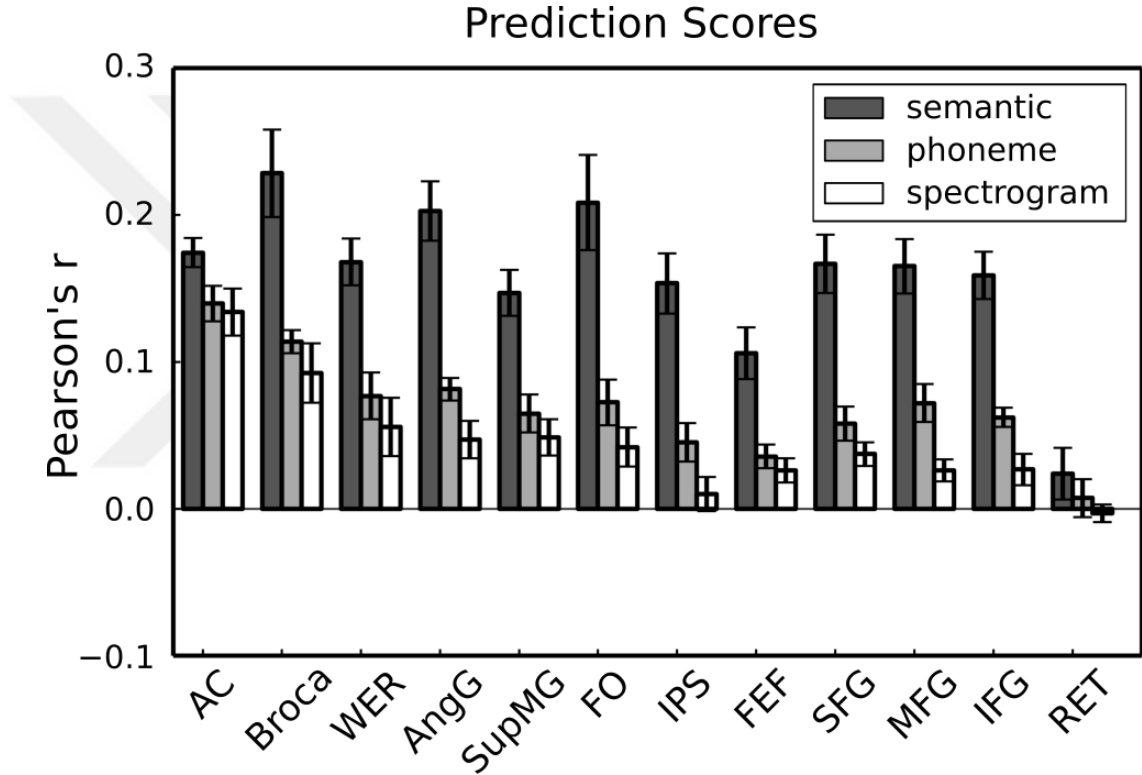


Figure A.6: **Prediction scores in functional ROIs.** Prediction scores are reported as mean  $\pm$  s.d. across subjects for the semantic model and the two control models, namely the spectrogram and phoneme models. In all functional ROIs, the average prediction score of the semantic model is significantly higher than that of the control models ( $P < 0.05$ , bootstrap test). Control models have relatively high prediction performances only in auditory cortex (AC). In high-level auditory and attention areas, the control models exhibit much poorer prediction performances than the semantic model.

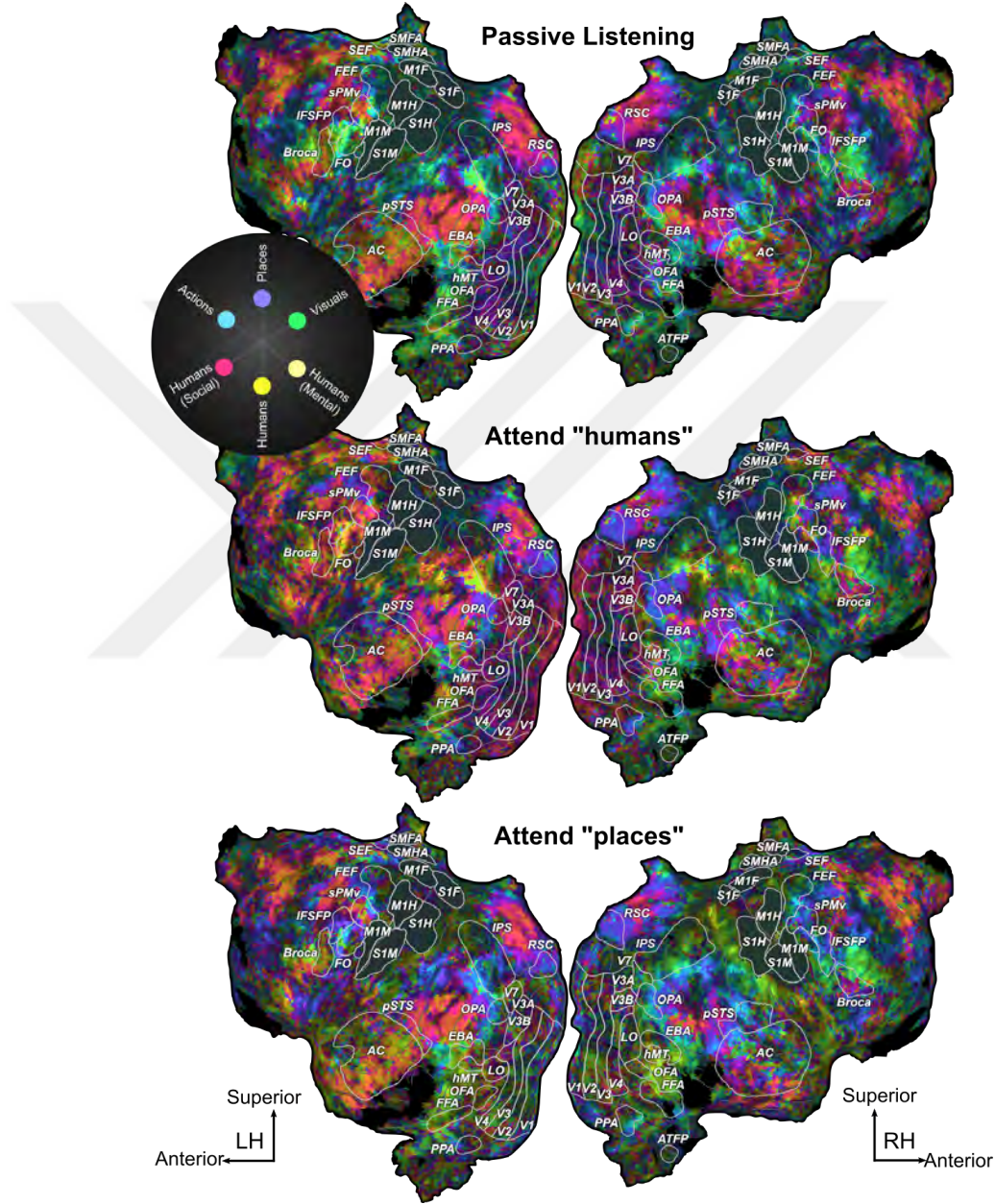


Figure A.7: **Cortical flat maps of semantic tuning for subject S1.** Distributions of semantic tuning are shown for passive listening and the two attention tasks. Each voxel's tuning vector was projected onto the three semantic dimensions, and RGB colors are assigned to each voxel according to these projections. Insignificant voxels are shown in gray. Yellow/green voxels are tuned to humans while blue voxels are tuned to places and locations. Results indicate that many cortical voxels shift their semantic tuning toward the attended word category.

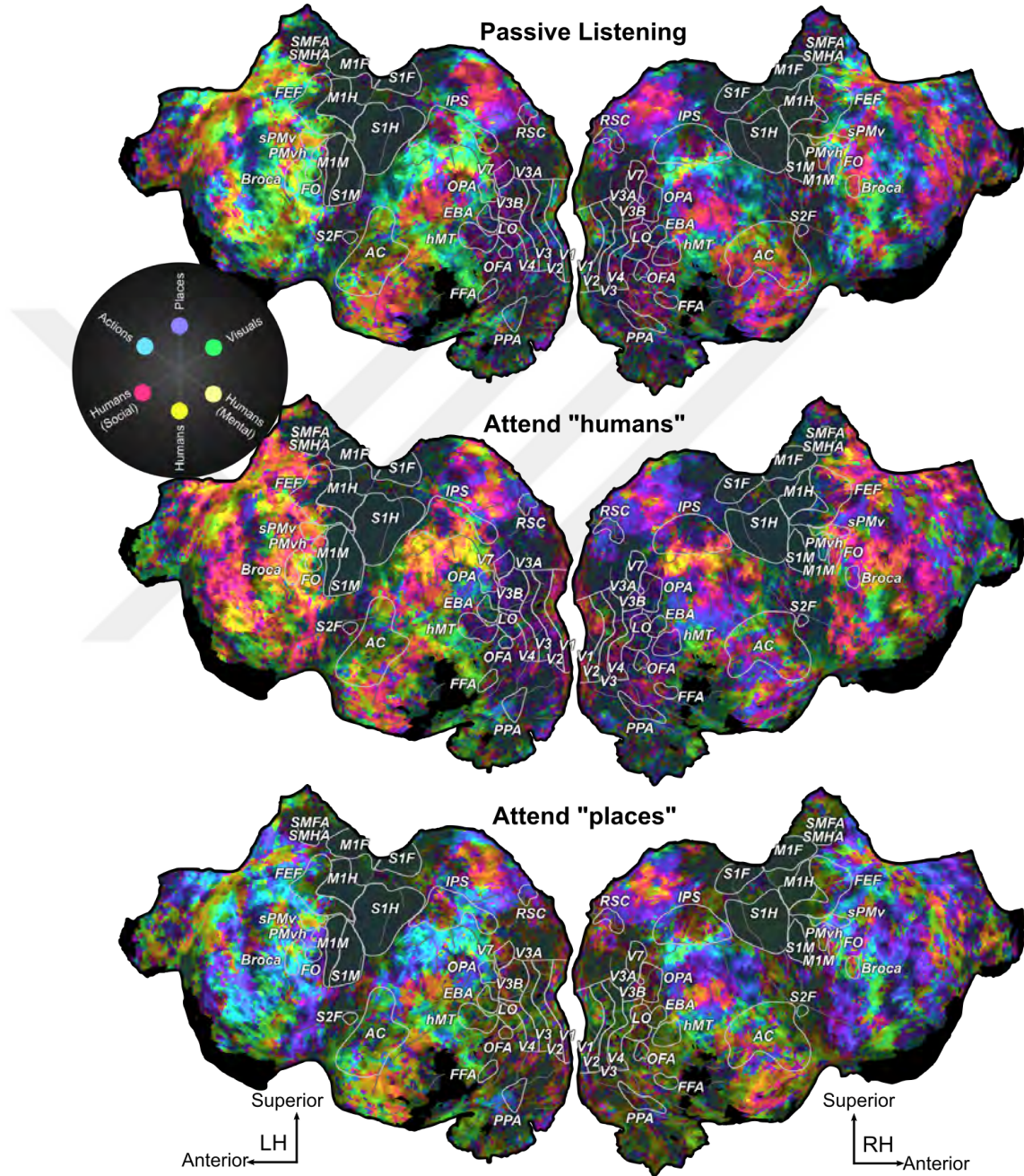


Figure A.8: **Cortical flat maps of semantic tuning for subject S2.** Distributions of semantic tuning are shown for passive listening and the two attention tasks. Each voxel's tuning vector was projected onto the three semantic dimensions, and RGB colors are assigned to each voxel according to these projections. Insignificant voxels are shown in gray. Yellow/green voxels are tuned to humans while blue voxels are tuned to places and locations. Results indicate that many cortical voxels shift their semantic tuning toward the attended word category.

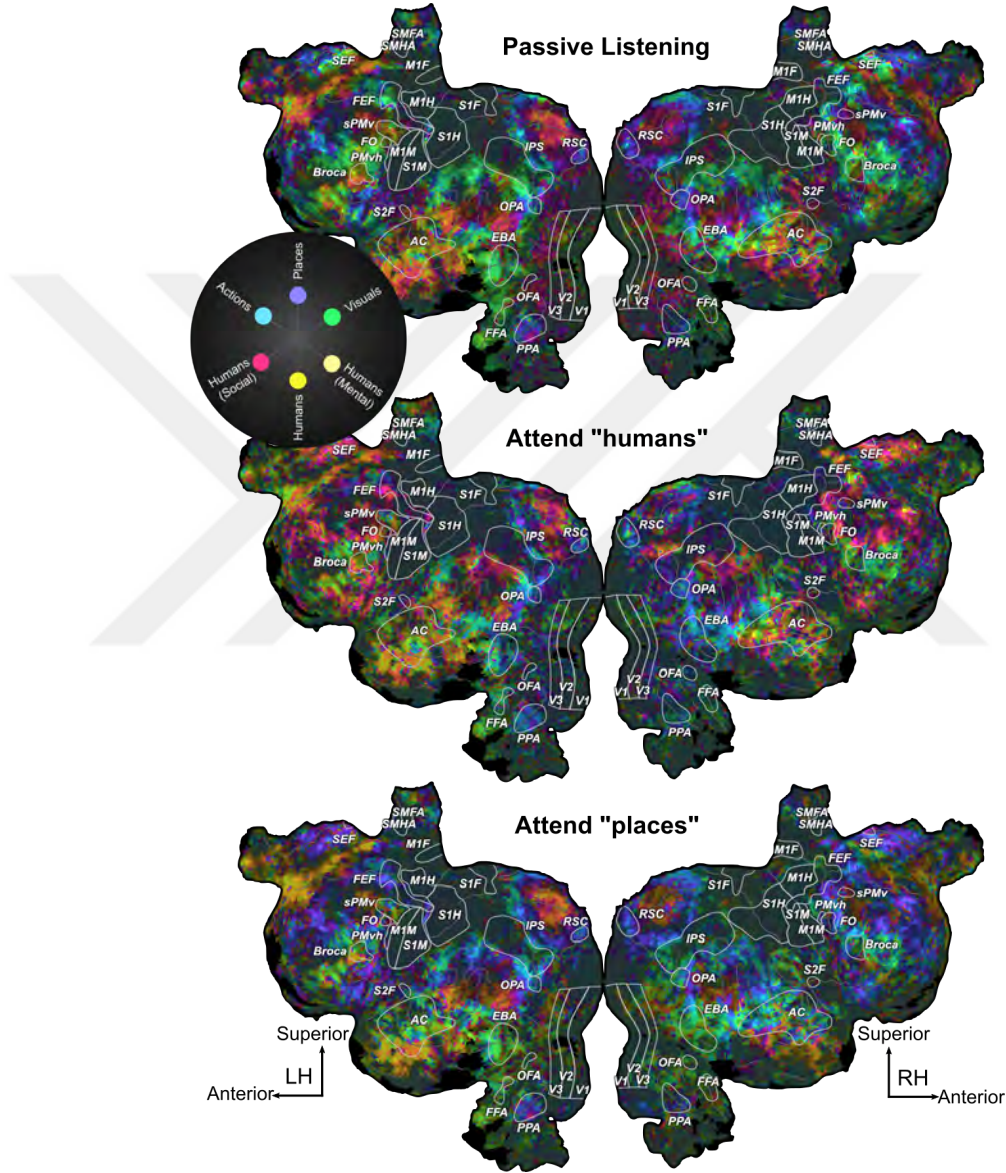


Figure A.9: **Cortical flat maps of semantic tuning for subject S3.** Distributions of semantic tuning are shown for passive listening and the two attention tasks. Each voxel's tuning vector was projected onto the three semantic dimensions, and RGB colors are assigned to each voxel according to these projections. Insignificant voxels are shown in gray. Yellow/green voxels are tuned to humans while blue voxels are tuned to places and locations. Results indicate that many cortical voxels shift their semantic tuning toward the attended word category.

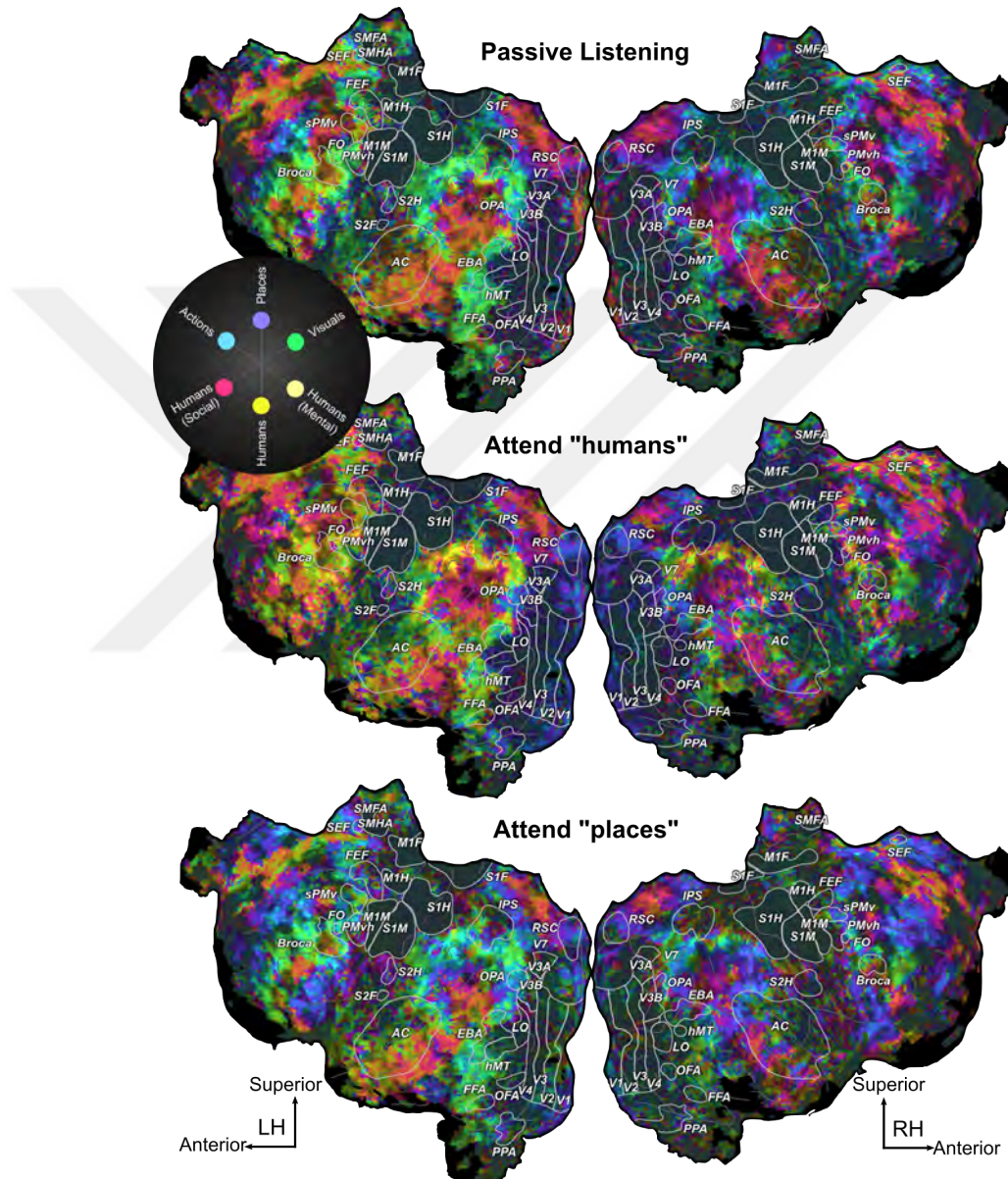


Figure A.10: **Cortical flat maps of semantic tuning for subject S4.** Distributions of semantic tuning are shown for passive listening and the two attention tasks. Each voxel's tuning vector was projected onto the three semantic dimensions, and RGB colors are assigned to each voxel according to these projections. Insignificant voxels are shown in gray. Yellow/green voxels are tuned to humans while blue voxels are tuned to places and locations. Results indicate that many cortical voxels shift their semantic tuning toward the attended word category.

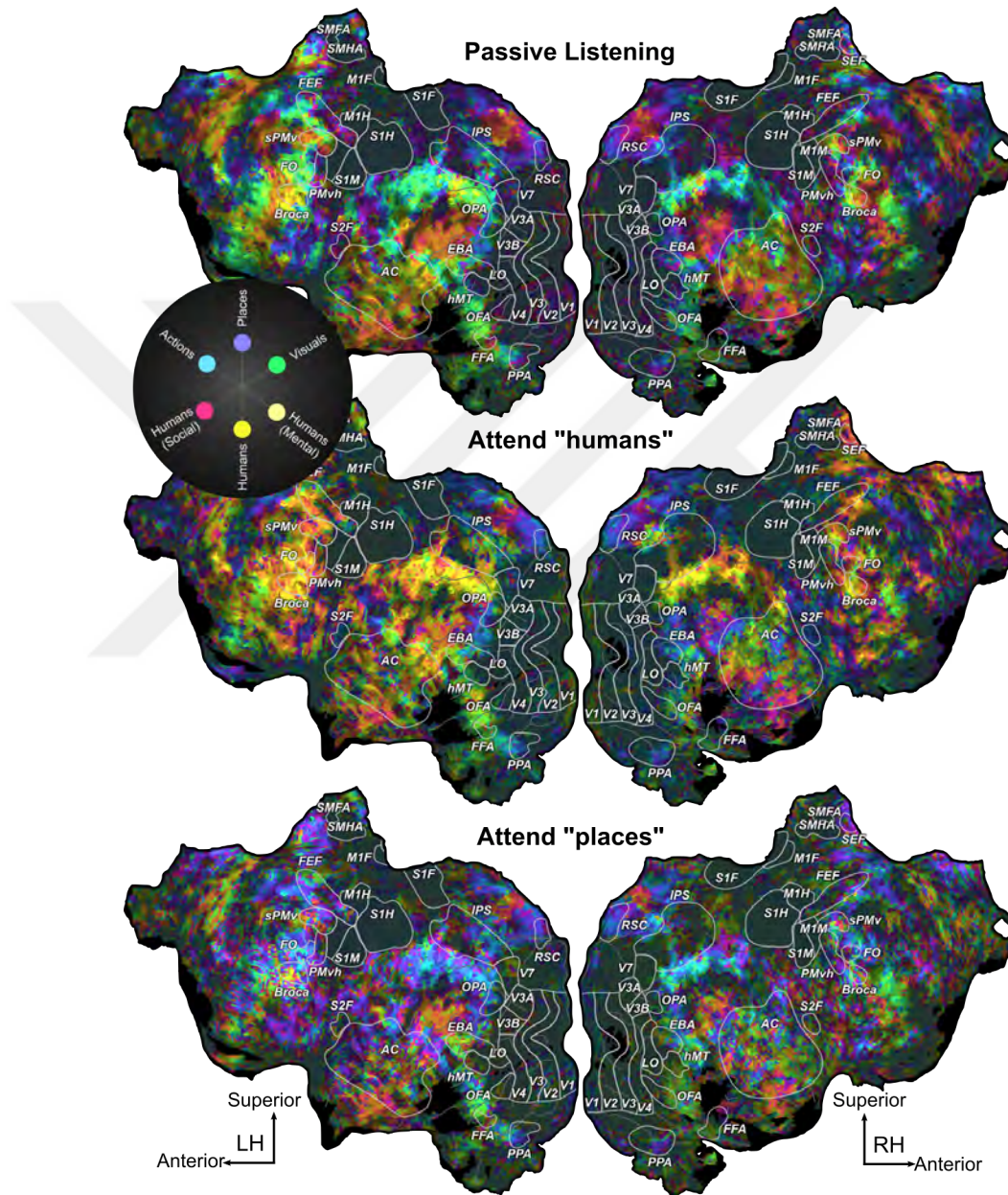


Figure A.11: **Cortical flat maps of semantic tuning for subject S5.** Distributions of semantic tuning are shown for passive listening and the two attention tasks. Each voxel's tuning vector was projected onto the three semantic dimensions, and RGB colors are assigned to each voxel according to these projections. Insignificant voxels are shown in gray. Yellow/green voxels are tuned to humans while blue voxels are tuned to places and locations. Results indicate that many cortical voxels shift their semantic tuning toward the attended word category.

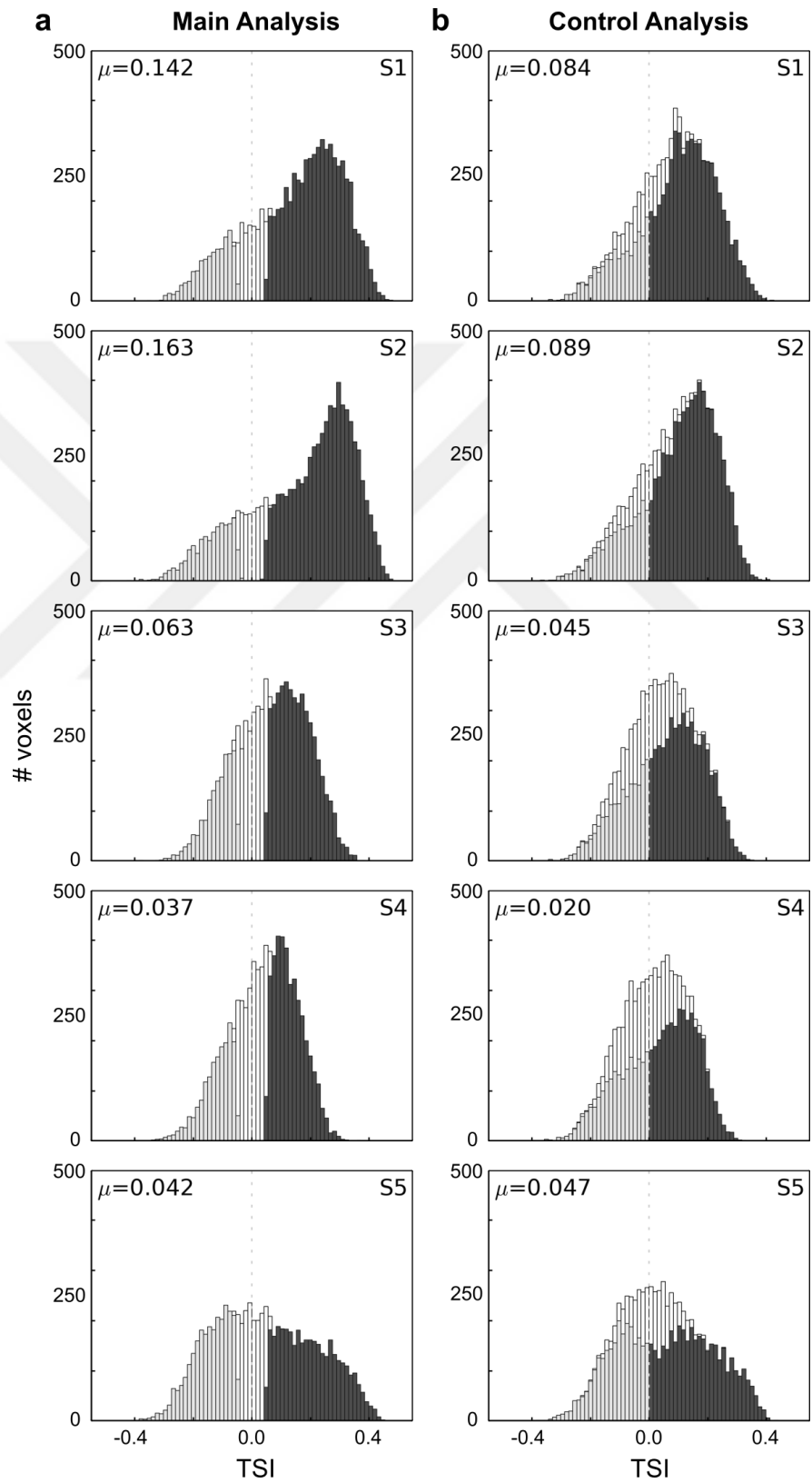


Figure A.12: Caption next page.

Figure A.12: **TSI histograms across cortical voxels.** (a) Main analysis. TSI distributions are shown for the cortical voxels (except motor areas) of subjects S1-S5. Only well-modeled (prediction score greater than 1 s.d. above the mean) and significantly attention-modulated ( $t$  test,  $q(FDR) < 0.05$ ) voxels are shown (7882-9014, depending on subject). Significantly positive TSIs (Wilcoxon signed-rank test,  $q(FDR) < 0.05$ ) are represented by dark gray bars and significantly negative TSIs (Wilcoxon signed-rank test,  $q(FDR) < 0.05$ ) are represented by light gray bars. Average TSI value is significantly greater than 0 in all subjects (Wilcoxon signed-rank test,  $P < 0.001$ ). These results indicate that attention shifts tuning of most well-modeled and attention-modulated cortical voxels toward the attended category. (b) Control analysis. To ensure that voxels shift their tuning towards attended categories even when targets are absent from the stimuli, a control analysis was done. Model fitting and TSI analyses were done using only those segments of the story stimuli that did not contain any words semantically related to humans or places. We then performed the TSI analyses on the fit models similarly to the main analysis. The TSI distributions in the absence of target stimuli are shown for the cortical voxels (except motor areas) of subjects S1-S5. Same voxel selection procedure is implemented as in (a). Average TSI value is significantly greater than 0 in all subjects (Wilcoxon signed-rank test,  $P < 0.001$ ). These results indicate that most well-modeled voxels shift their tuning toward the target category even when the target is not present in the stimuli.

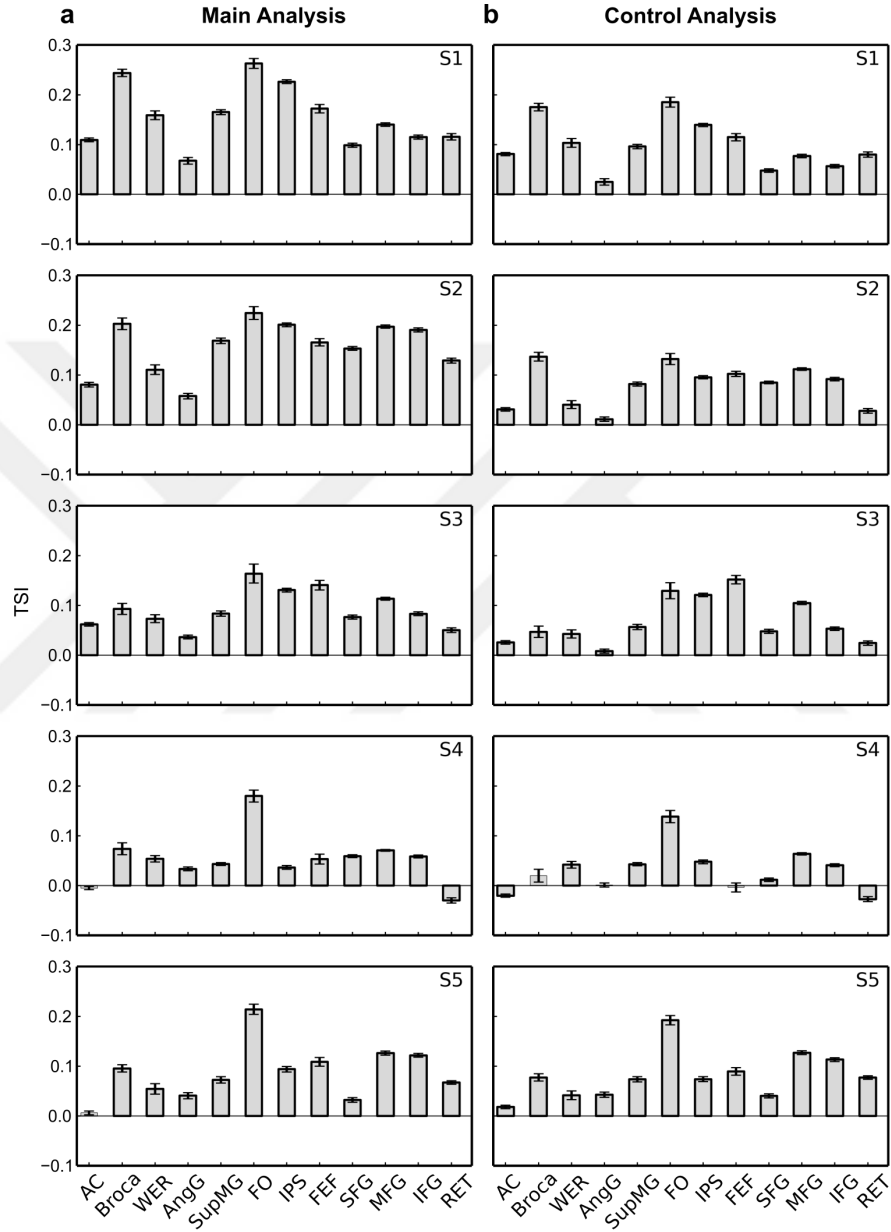


Figure A.13: **Distributions of tuning shifts across functional ROIs.** (a) Main analysis. Average TSIs of functional ROIs are shown for subjects S1-S5 (mean  $\pm$  s.e.m. across voxels). Results indicate that, attention causes different levels of semantic tuning shifts across functional ROIs. Areas that take place in attention control (FO, IPS, FEF), and language processing (Broca, WER) have high TSI for all subjects. (b) Control analysis. TSIs in functional ROIs are shown for subjects S1-S5 in the absence of targets in the stimuli (mean  $\pm$  s.e.m. across voxels). Results indicate that even when no target is present in the stimuli, functional ROIs have TSI characteristics similar to main analysis.

## Appendix B

### Supplementary Figures-2

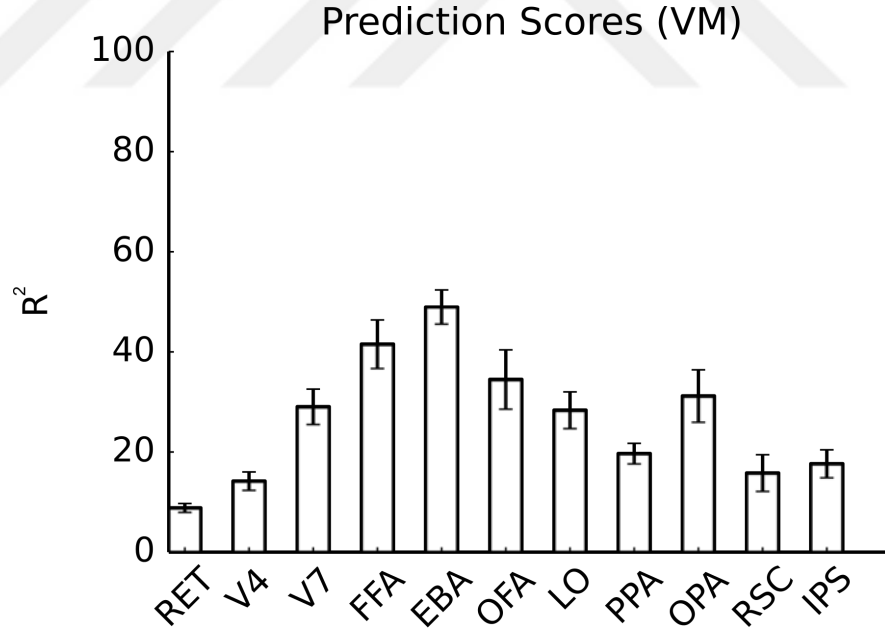


Figure B.1: **Prediction scores of VM across functional ROIs.** The prediction performance of VM is shown for several well-known functional ROIs. Bar graphs display improvement in prediction scores ( $R^2$ , mean  $\pm$  std across subjects). VM has significantly positive prediction scores in all ROIs ( $P < 0.05$ , Wilcoxon signed rank test). The average prediction score is  $33.32 \pm 2.98$  in category-selective areas and  $10.37 \pm 1.54$  in early visual areas.

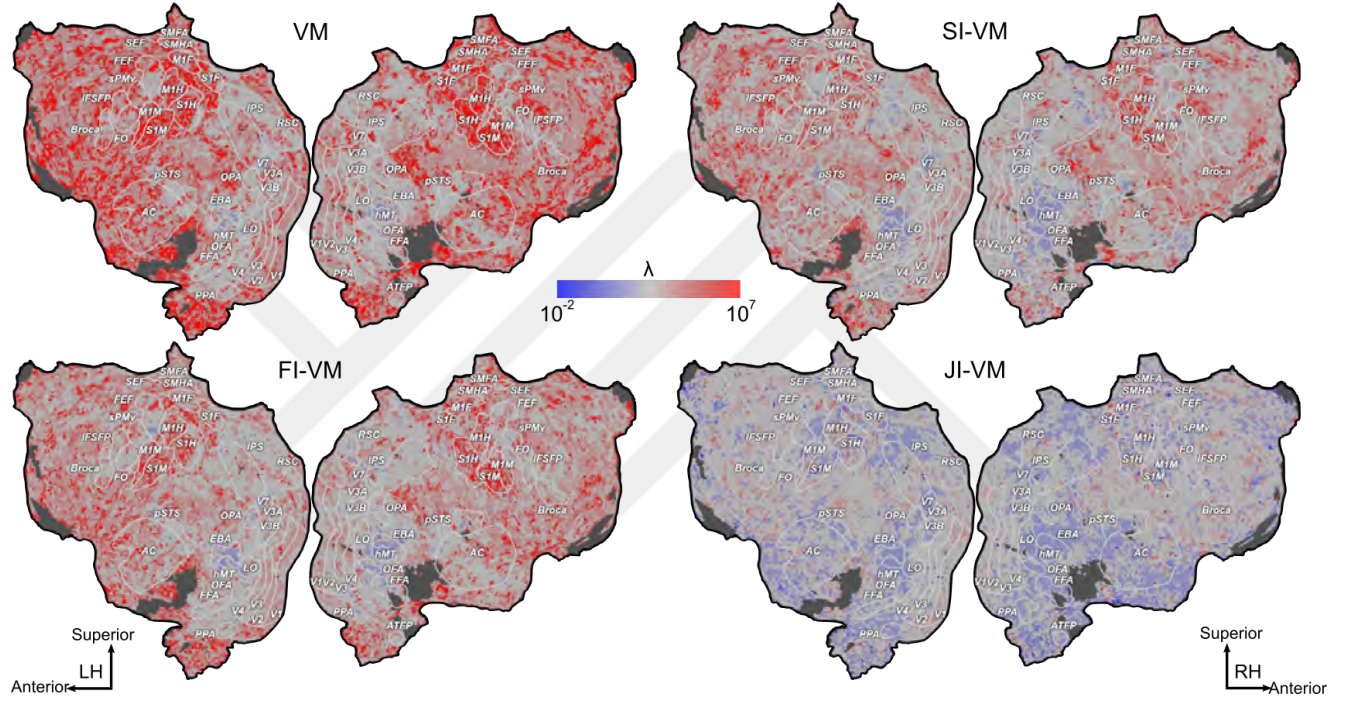


Figure B.2: **Cortical distribution of optimal L2-norm regularization parameter for Subject S1.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity.

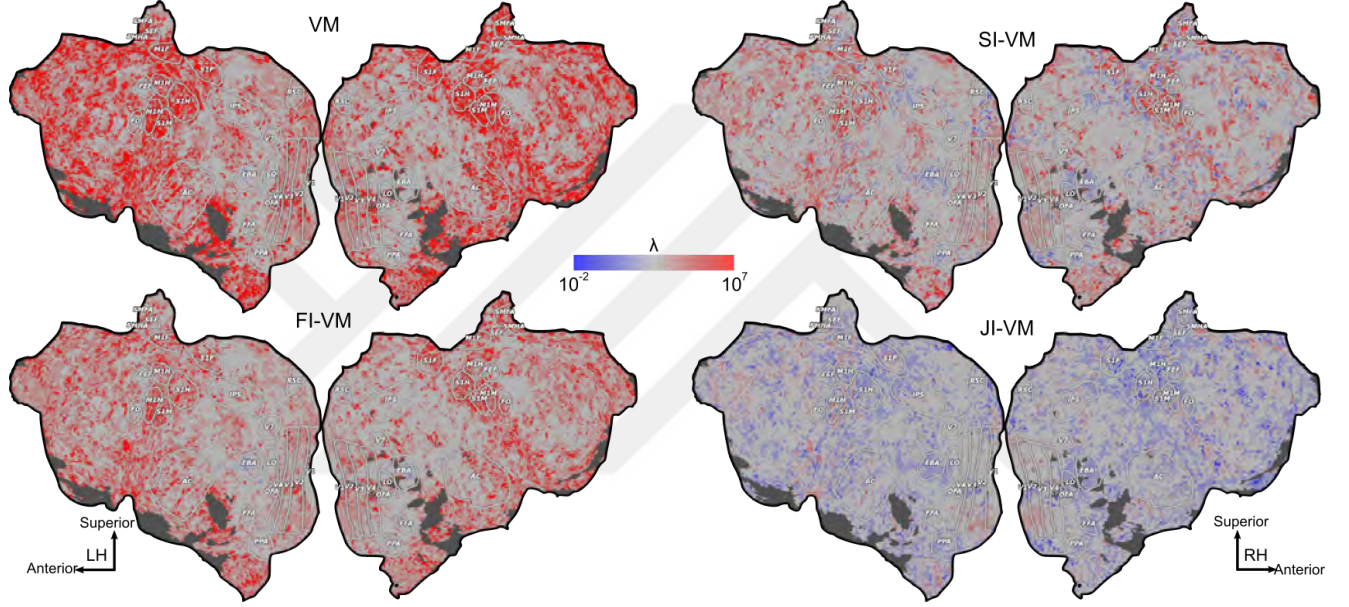


Figure B.3: **Cortical distribution of optimal L2-norm regularization parameter for Subject S2.** The  $L_2$ -regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by  $L_2$ -regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for  $L_2$ -regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity.

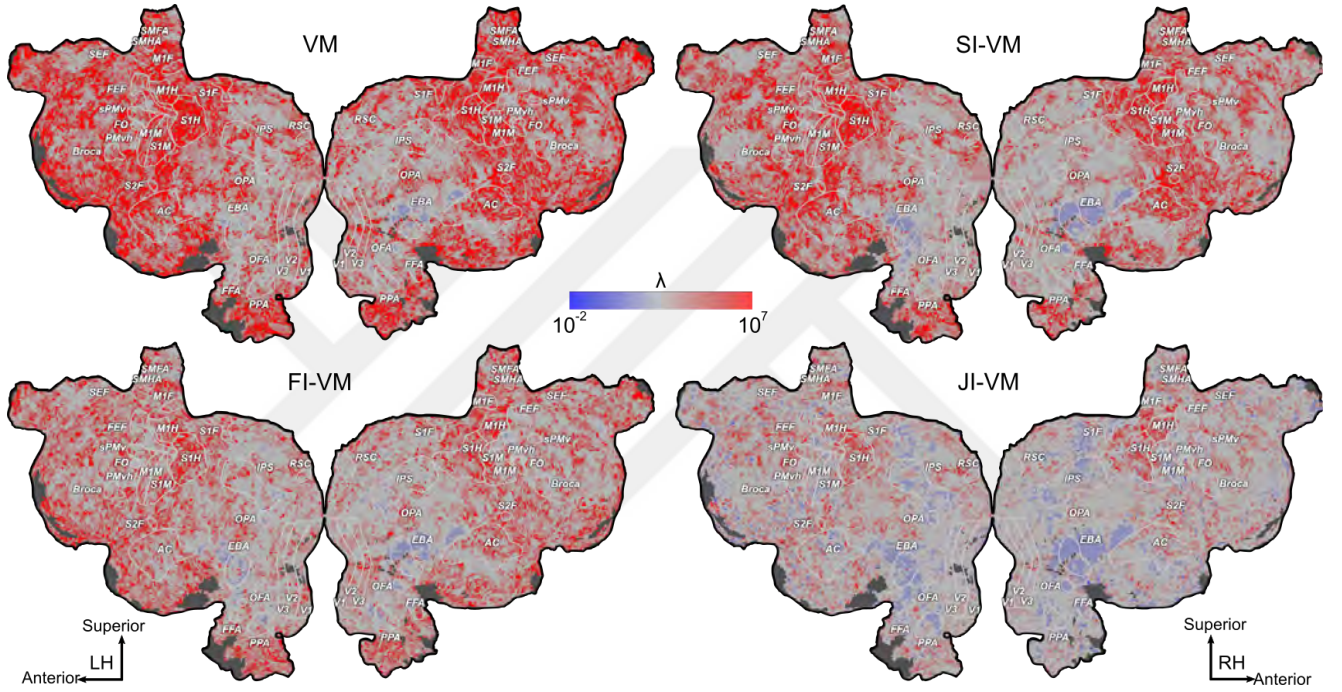


Figure B.4: **Cortical distribution of optimal L<sub>2</sub>-norm regularization parameter for Subject S3.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity.

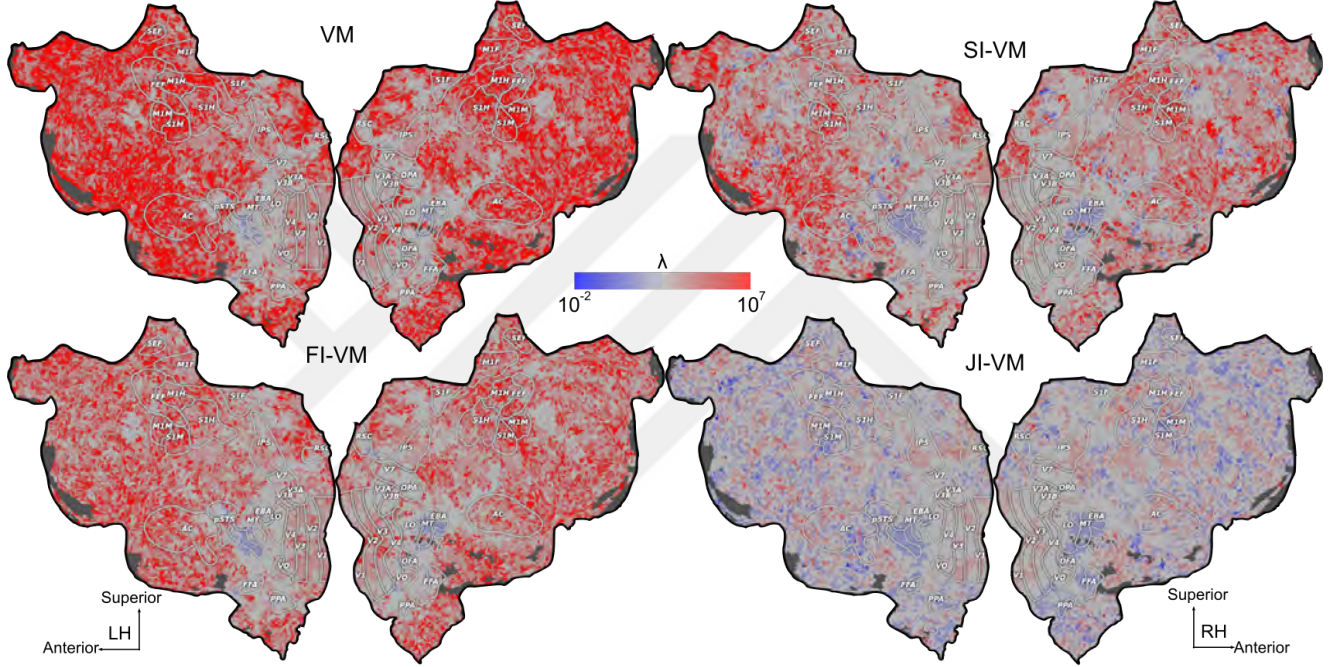


Figure B.5: **Cortical distribution of optimal L2-norm regularization parameter for Subject S4.** The  $L_2$ -regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by  $L_2$ -regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for  $L_2$ -regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity.

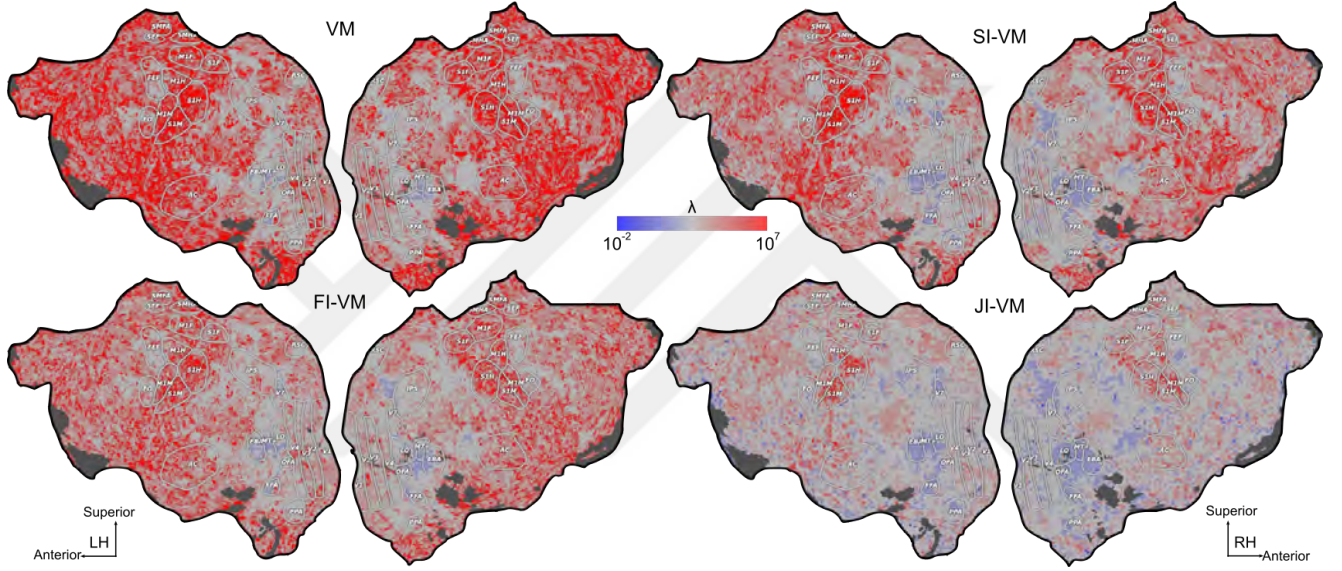


Figure B.6: **Cortical distribution of optimal L2-norm regularization parameter for Subject S5.** The L<sub>2</sub>-regularization parameters over model features in three methods (VM, FI-VM, JI-VM) are visualized on flat maps. VM strictly penalizes model weights broadly across cortex. High regularization parameters reduce sensitivity in measuring voxelwise functional selectivity. Here, we incorporate additional regularization terms across features and across voxel neighborhoods. As such, the proposed method should limit unnecessary penalization of model weights by L<sub>2</sub>-regularization. As a result, FI-VM applies less weight regularization particularly in category-selective areas, early visual areas, and frontal gyrus. Overall, JI-VM enforces the lowest weight penalization among the competing methods consistently across cortex. These results imply that both spatial and feature regularization alleviate the need for L<sub>2</sub>-regularization, thereby increasing sensitivity in measuring voxelwise functional selectivity.

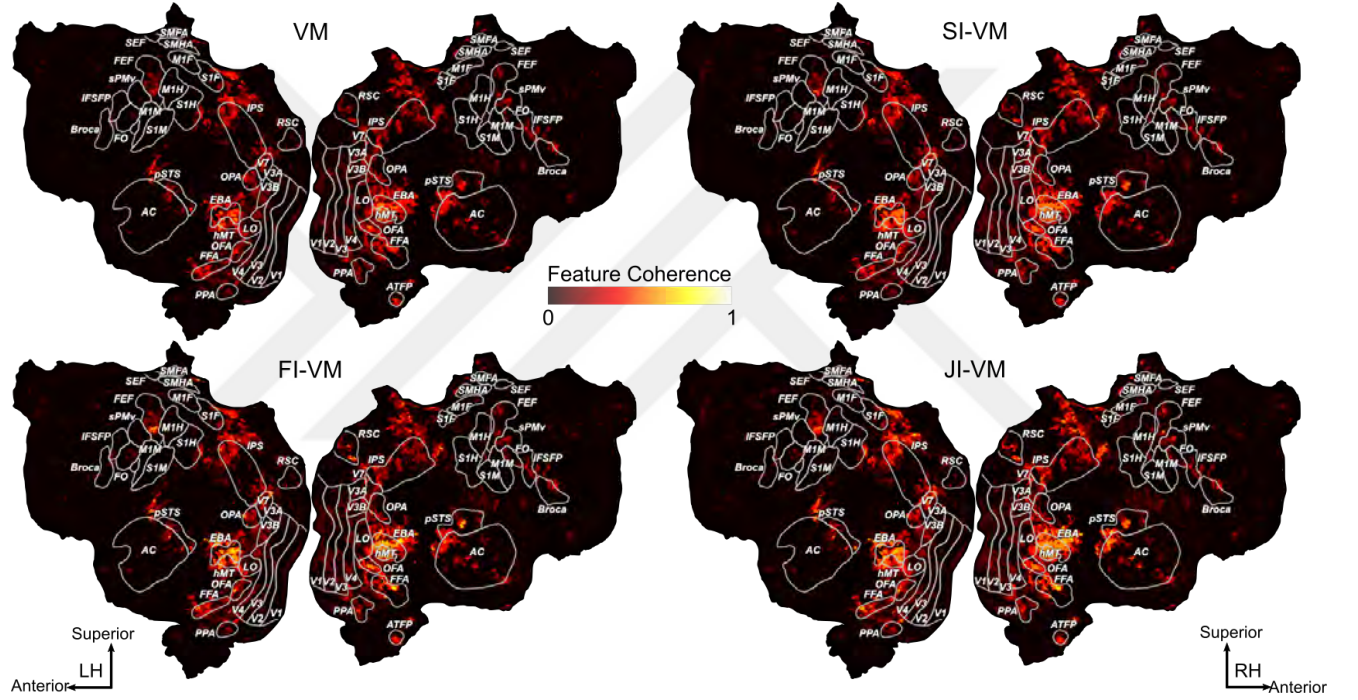


Figure B.7: **Feature coherence in VM, FI-VM, and JI-VM for Subject S1.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and AITFP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex.

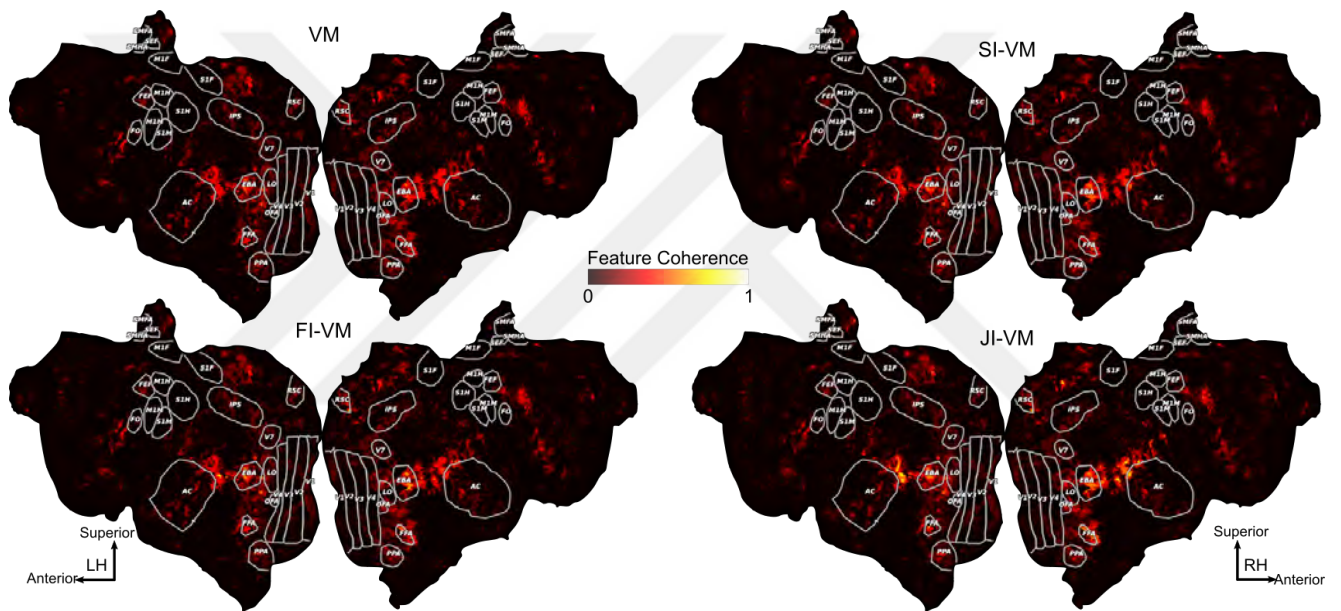


Figure B.8: **Feature coherence in VM, FI-VM, and JI-VM for Subject S2.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and ATRP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex.

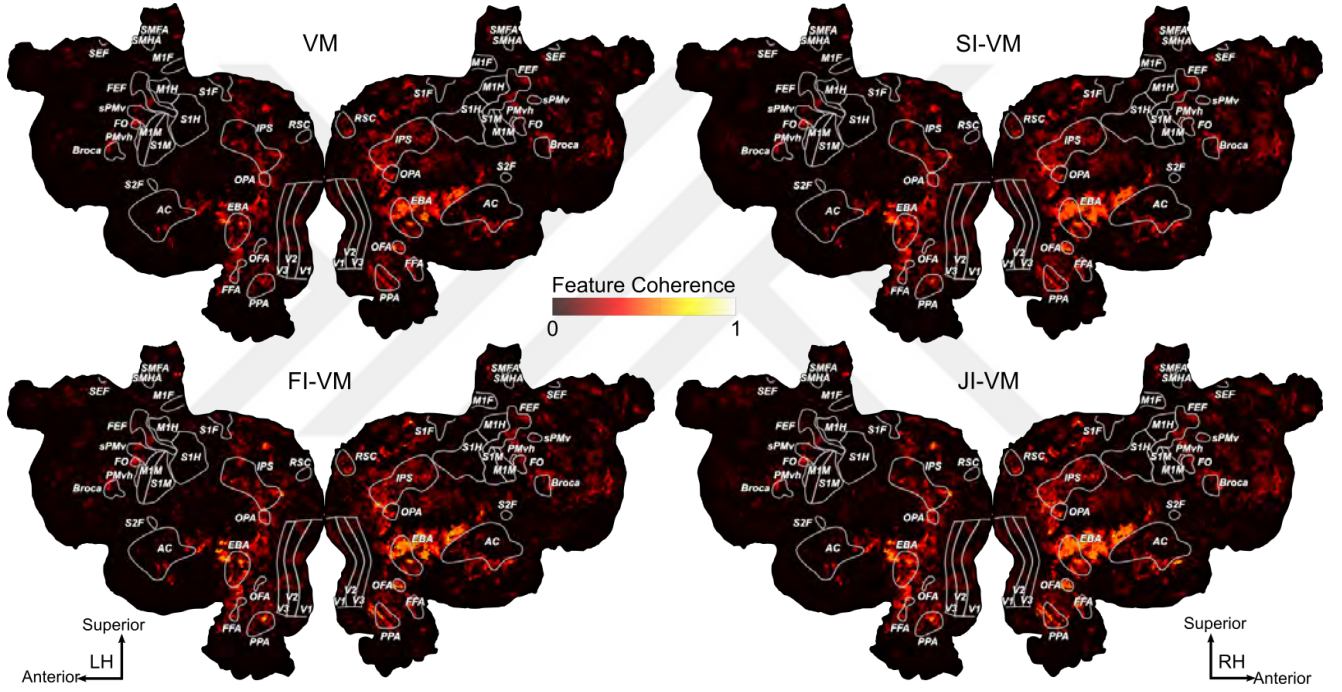


Figure B.9: **Feature coherence in VM, FI-VM, and JI-VM for Subject S3.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and ATRP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex.

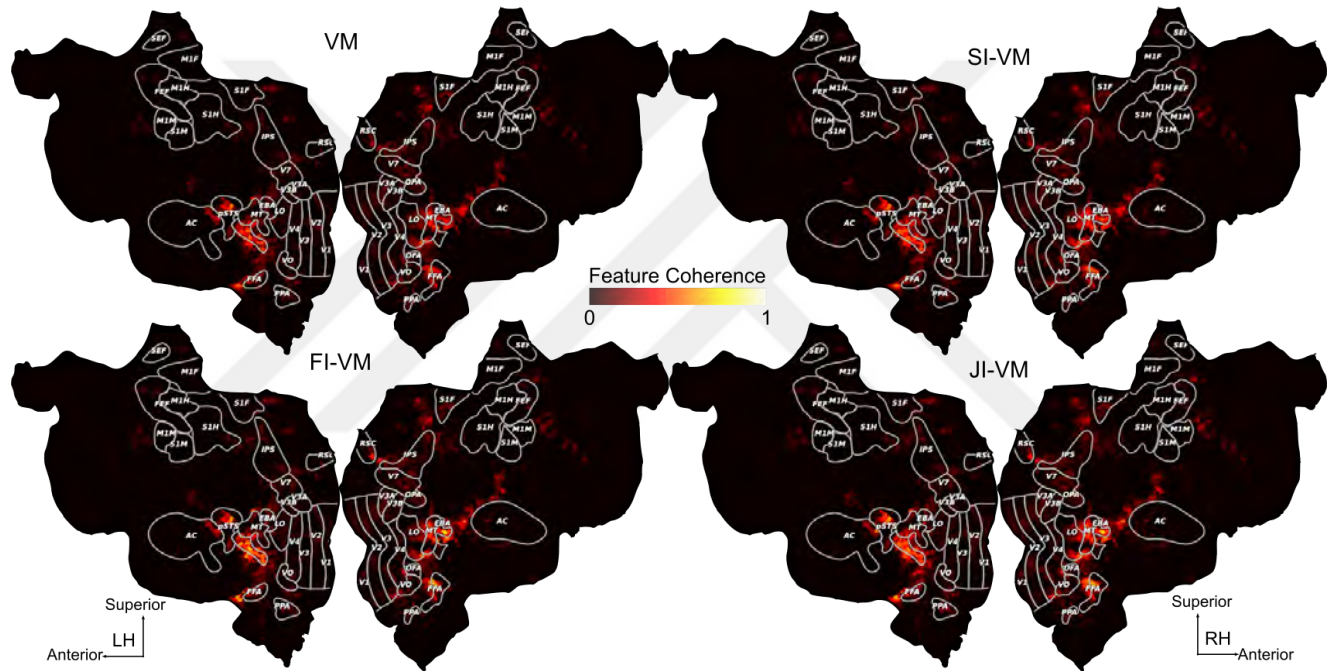


Figure B.10: **Feature coherence in VM, FI-VM, and JI-VM for Subject S4.** Feature coherence is shown on flat maps. VM ignores potential feature correlations in the stimulus so it has reduced sensitivity to coherence in stimulus features. The informed feature regularization enforces similar model weights for correlated model features. Therefore, feature coherence of FI-VM and JI-VM are higher than that of VM across many cortical regions including category-selective areas, early visual areas, RSC, IPS, and ATRP. These results suggest that informed feature regularization improves feature coherence broadly across visual cortex.



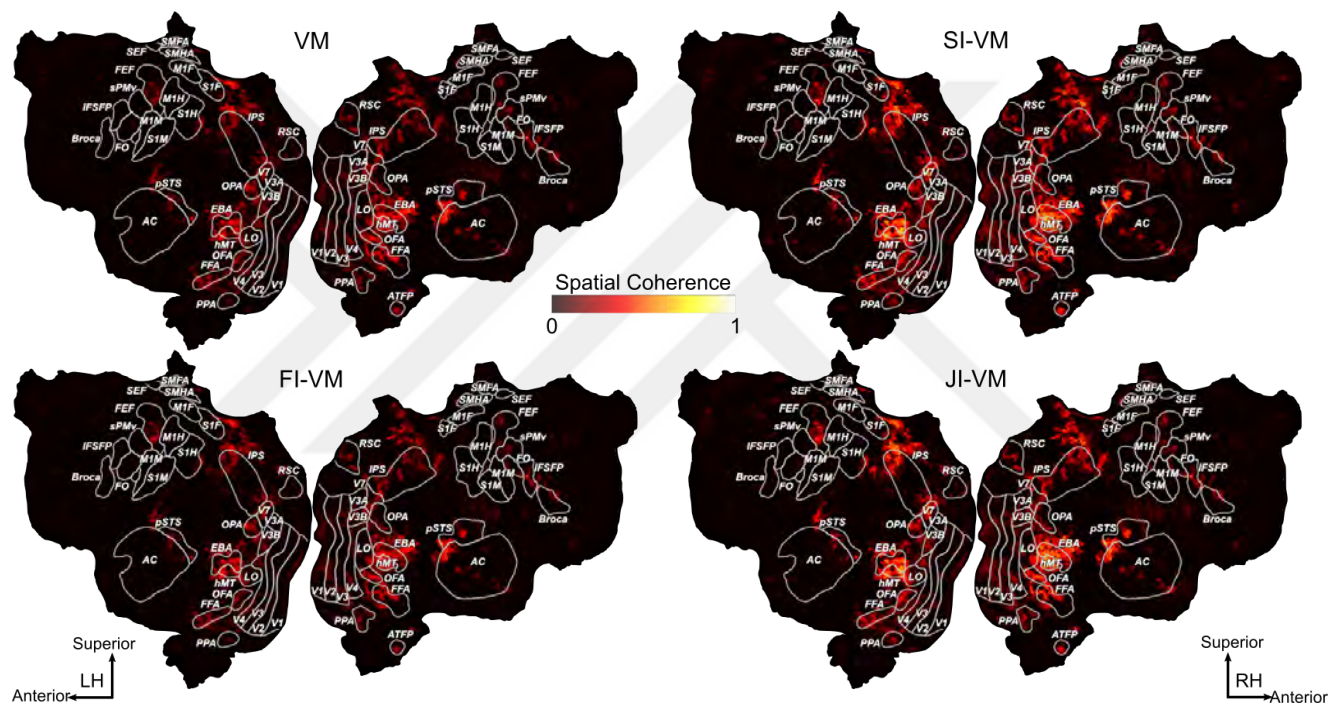


Figure B.12: **Spatial coherence in VM, FI-VM, and JI-VM for Subject S1.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and ATEP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex.

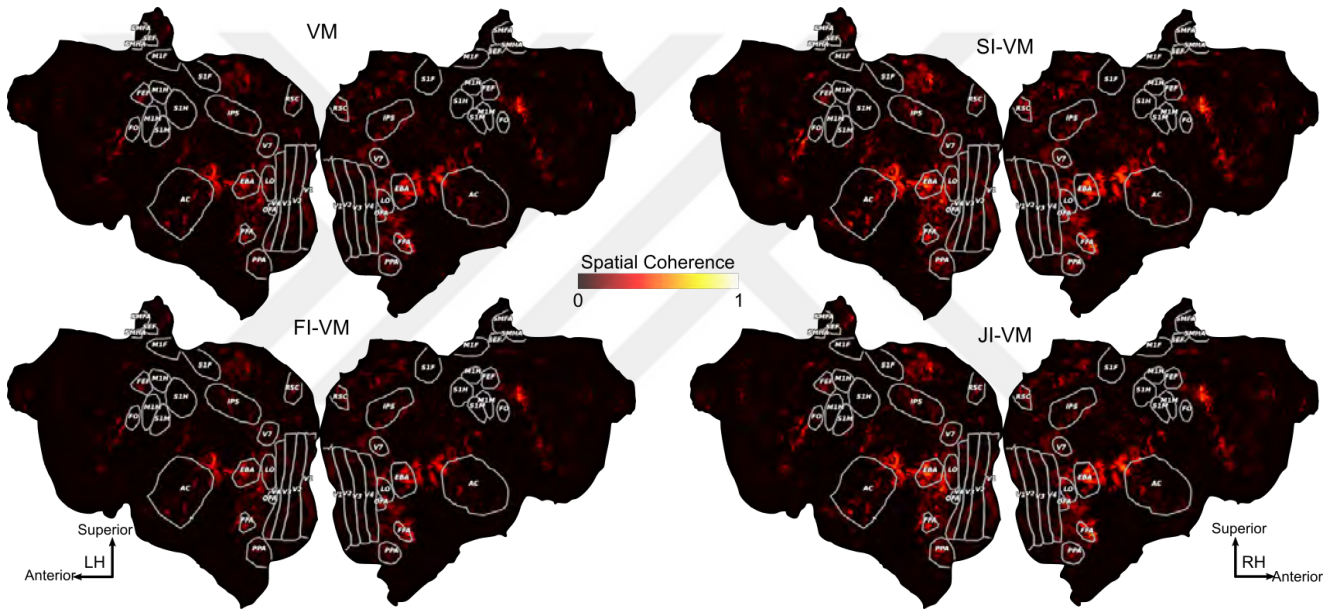


Figure B.13: **Spatial coherence in VM, FI-VM, and JI-VM for Subject S2.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITFP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex.



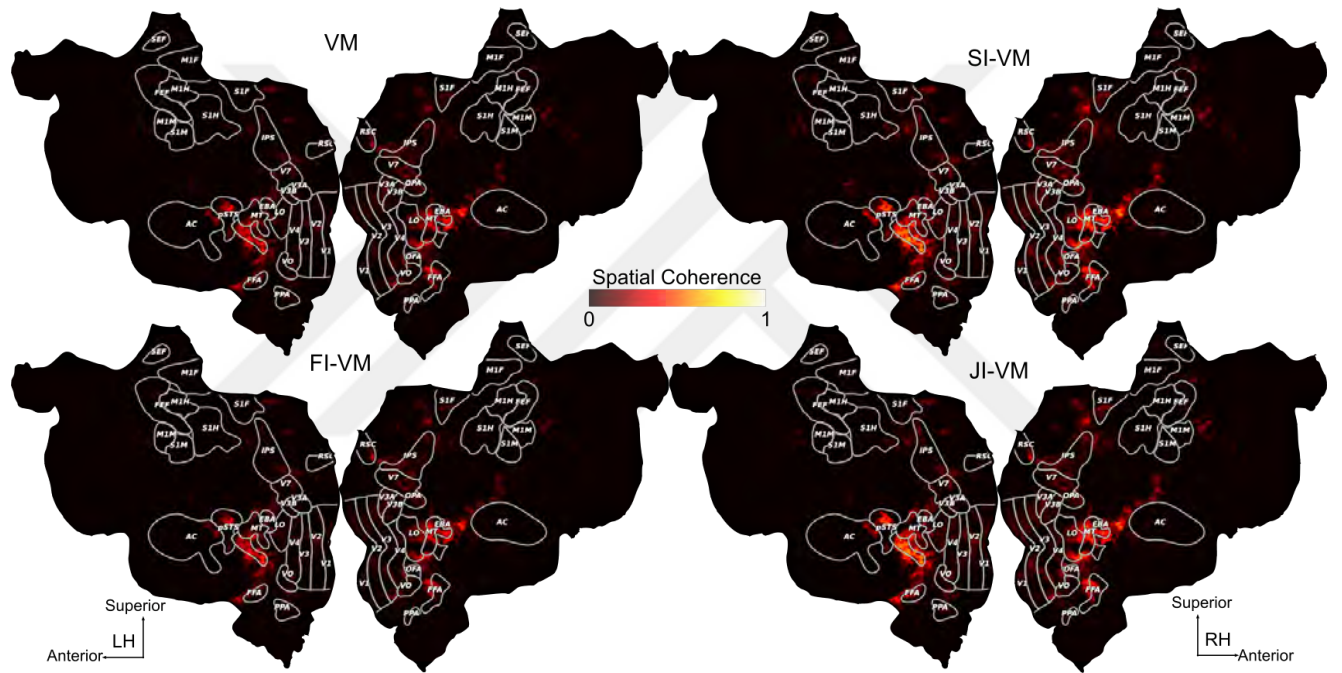


Figure B.15: **Spatial coherence in VM, FI-VM, and JI-VM for Subject S4.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITFP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex.

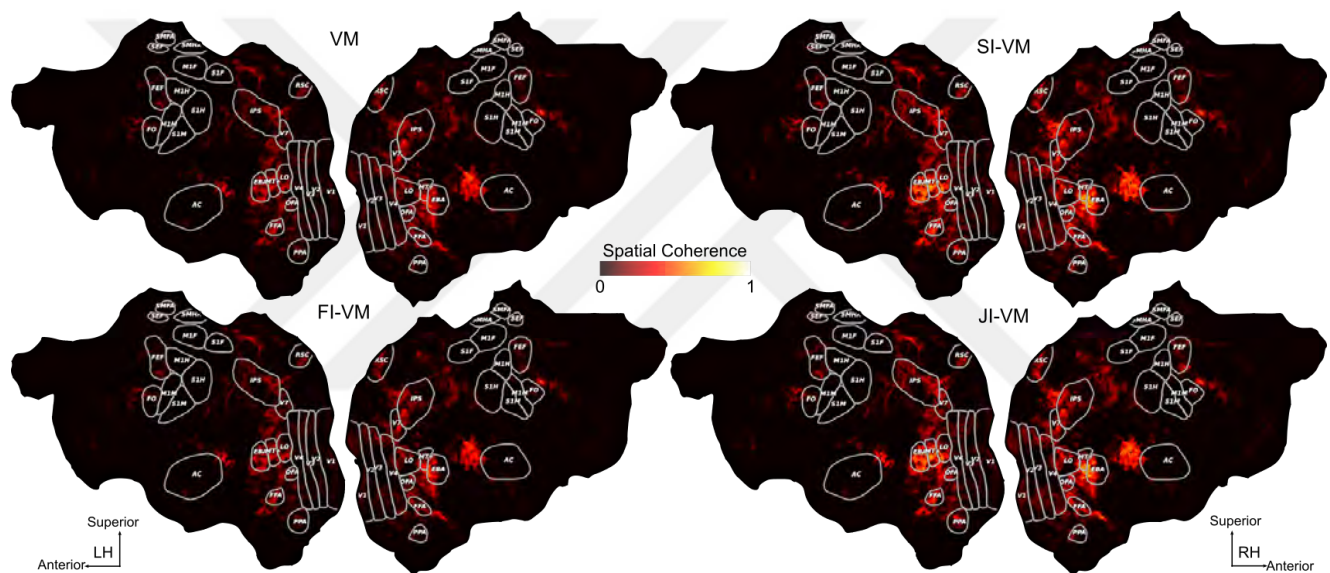


Figure B.16: **Spatial coherence in VM, FI-VM, and JI-VM for Subject S5.** Spatial coherence is shown on flat maps. VM ignores correlations among voxel neighborhoods so it has reduced sensitivity to spatial coherence in cortical representations. Spatial regularization enforces increased spatial coherence among cortical representations within voxel neighborhoods. Although, FI-VM and VM have relatively similar spatial coherence values, JI-VM consistently yields higher spatial coherence across many cortical regions including category-selective areas, early visual areas (V1-V4), IPS, RSC and AITFP. These results suggest that spatial regularization improves spatial coherence broadly across visual cortex.