ARTICLE

doi:10.1038/nature17637

Natural speech reveals the semantic maps that tile human cerebral cortex

Alexander G. Huth[1], Wendy A. de Heer[2], Thomas L. Griffiths[1,2], Frédéric E. Theunissen[1,2] & Jack L. Gallant[1,2]

The meaning of language is represented in regions of the cerebral cortex collectively known as the 'semantic system'. However, little of the semantic system has been mapped comprehensively, and the semantic selectivity of most regions is unknown. Here we systematically map semantic selectivity across the cortex using voxel-wise modelling of functional MRI (fMRI) data collected while subjects listened to hours of narrative stories. We show that the semantic system is organized into intricate patterns that seem to be consistent across individuals. We then use a novel generative model to create a detailed semantic atlas. Our results suggest that most areas within the semantic system represent information about specific semantic domains, or groups of related concepts, and our atlas shows which domains are represented in each area. This study demonstrates that data-driven methods—commonplace in studies of human neuroanatomy and functional connectivity—provide a powerful and efficient means for mapping functional representations in the brain.

Previous neuroimaging studies have identified a group of regions that seem to represent information about the meaning of language. These regions, collectively known as the semantic system, respond more to words than non-words[1], more to semantic tasks than phonological tasks[1], and more to natural speech than temporally scrambled speech[2]. Studies that have investigated specific types of representation in the semantic system have found areas selective for concrete or abstract words[3–5], action verbs[6], social narratives[7] or other semantic features. Others have found areas selective for specific semantic domains—groups of related concepts such as living things, tools, food or shelter[8–13]. However, all previous studies tested only a handful of stimulus conditions, so no study has yet produced a comprehensive survey of how semantic information is represented across the entire semantic system.

We addressed this problem by using a data-driven approach[14] to model brain responses elicited by naturally spoken narrative stories that contain many different semantic domains[15]. Seven subjects listened to more than two hours of stories from *The Moth Radio Hour*[2] while whole-brain blood-oxygen-level-dependent (BOLD) responses were recorded by fMRI. We then used voxel-wise modelling, a highly effective approach for modelling responses to complex natural stimuli[14–17], to estimate the semantic selectivity of each voxel (Fig. 1a).

Voxel-wise model estimation and validation

In voxel-wise modelling, features of interest are first extracted from the stimuli and then regression is used to determine how each feature modulates BOLD responses in each voxel. We used a word embedding space to identify semantic features of each word in the stories[12,15,18–20]. The embedding space was constructed by computing the normalized co-occurrence between each word and a set of 985 common English words (such as 'above', 'worry' and 'mother') across a large corpus of English text. Words related to the same semantic domain tend to occur in similar contexts, and so have similar co-occurrence values. For example, the words 'month' and 'week' are very similar (the correlation between the two is 0.74), while the words 'month' and 'tall' are not (correlation −0.22).

Next we used regularized linear regression to estimate how the 985 semantic features influenced BOLD responses in every cortical voxel and in each individual subject (Fig. 1a). To account for responses caused by low-level properties of the stimulus such as word rate and phonemic content, additional regressors were included during voxel-wise model estimation and then discarded before further analysis. We also included additional regressors to account for physiological and emotional factors, but these had no effect on the estimated semantic models (Supplementary Data 3).

One advantage of voxel-wise modelling over conventional neuroimaging approaches is that the fit models can be validated by predicting BOLD responses to new natural stimuli that were not used during model estimation. This makes it possible to compute effect size by finding the fraction of response variance explained by the models. We tested how well the voxel-wise models predicted BOLD responses elicited by a new 10-min *Moth* story (Fig. 1b) that had not been used for model estimation. We found good prediction performance for voxels located throughout the semantic system, including in the lateral temporal cortex (LTC) and ventral temporal cortex (VTC), lateral parietal cortex (LPC) and medial parietal cortex (MPC), and medial prefrontal cortex, superior prefrontal cortex (SPFC) and inferior prefrontal cortex (IPFC) (Fig. 1c and Extended Data Fig. 1). This suggests that much of the semantic system is domain selective.

Mapping semantic representation across cortex

By inspecting the fit models, we can determine which specific semantic domains are represented in each voxel. In theory this could be done by examining each voxel separately. However, our data consist of tens of thousands of voxels per subject, rendering this approach unfeasible. A practical alternative is to project the models into a low-dimensional subspace that retains as much information as possible about the semantic tuning of the voxels[10,14]. We found such a space by applying principal components analysis to the estimated models aggregated across subjects, producing 985 orthogonal semantic dimensions that are ordered by how much variance each explained across the voxels. It is likely that only some of these dimensions capture shared aspects of semantic tuning across the subjects; the rest reflect individual differences, fMRI noise, or the statistical properties of the stories. To identify the shared dimensions, we tested whether each explained more variance across the models than expected by chance, which was defined by the principal components of the stimulus matrix used for model estimation[14]. At least four dimensions explained a significant amount

[1]Helen Wills Neuroscience Institute, University of California, Berkeley, California 94720, USA. [2]Department of Psychology, University of California, Berkeley, California 94720, USA.

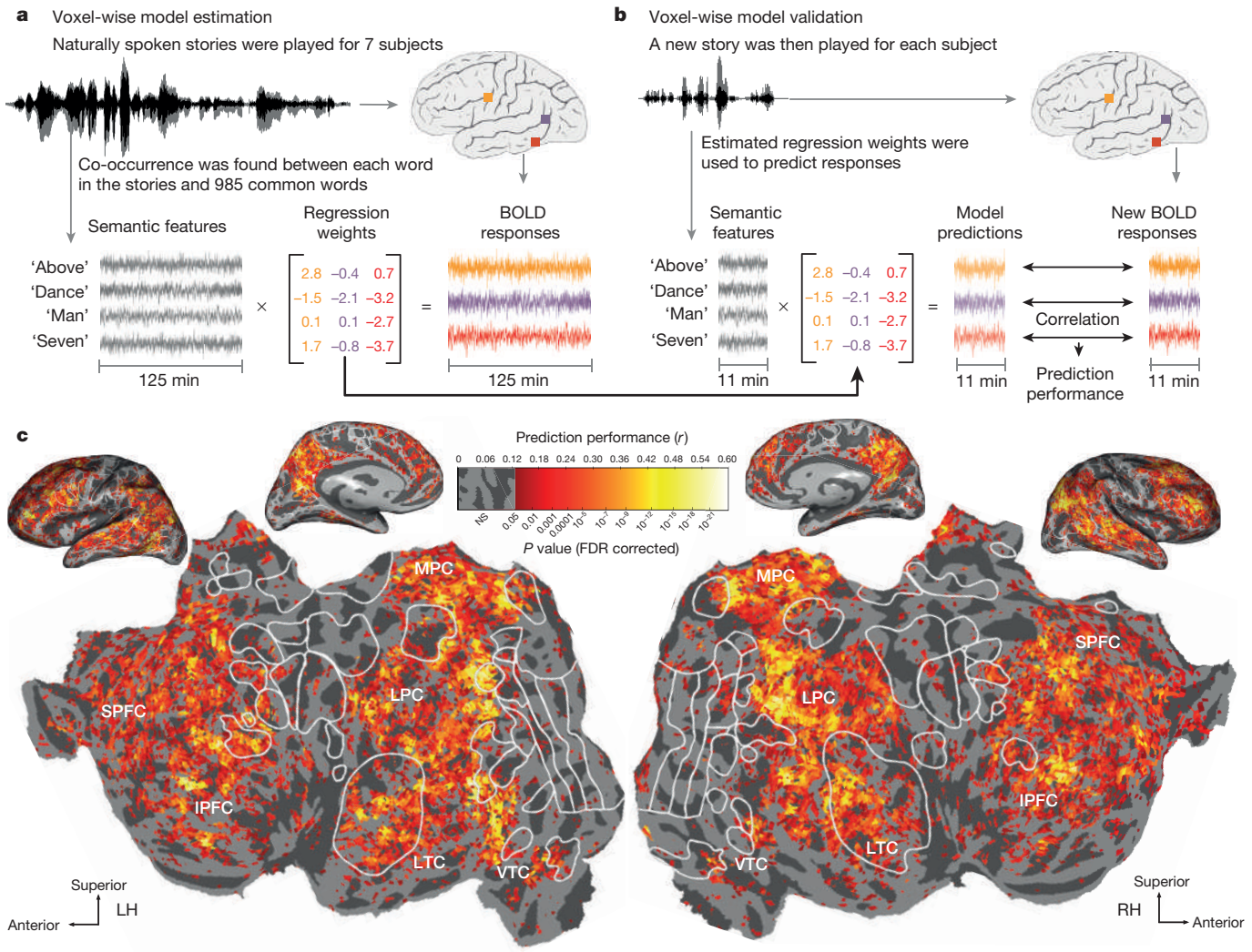


Figure 1 | Voxel-wise modelling. a, Seven subjects listened to over 2 h of naturally spoken narrative stories while BOLD responses were measured using fMRI. Each word in the stories was projected into a 985-dimensional word embedding space constructed using word co-occurrence statistics from a large corpus of text. A finite impulse response (FIR) regression model was estimated individually for every voxel. The voxel-wise model weights describe how words appearing in the stories influence BOLD signals. **b**, Models were tested using one 10-min story that was not included during model estimation. Model prediction performance was computed as the correlation between predicted responses to this story and actual BOLD responses. **c**, Prediction performance of voxel-wise models for one subject. Semantic models accurately predict BOLD responses in many brain areas, including the LTC, VTC, LPC, MPC, SPFC and IPFC. These regions have previously been identified as the semantic system in the human brain. LH, left hemisphere; RH, right hemisphere.

of variance ($P < 0.001$, Bonferroni-corrected bootstrap test) in all but one subject; in the last subject only three dimensions were significant (Extended Data Fig. 2). This suggests that our fMRI data contain about four statistically significant semantic dimensions that are shared across subjects.

The four shared semantic dimensions provide a way to summarize succinctly the semantic selectivity of a voxel. However, to interpret projections of the models onto these dimensions we need to understand how semantic information is encoded in this four-dimensional space. To visualize the semantic space, we projected the 10,470 words in the stories from the word embedding space onto each dimension. We then used *k*-means clustering to identify 12 distinct categories (see Supplementary Methods for details). Each category was inspected and labelled by hand. The labels assigned to the 12 categories were 'tactile' (a cluster containing words such as 'fingers'), 'visual' (words such as 'yellow'), 'numeric' ('four'), 'locational' ('stadium'), 'abstract' ('natural'), 'temporal' ('minute'), 'professional' ('meetings'), 'violent' ('lethal'), 'communal' ('schools'), 'mental' ('asleep'), 'emotional' ('despised') and 'social' ('child'). (See Supplementary Table 2 and Supplementary Data 5 for more detailed evaluations of each category.)

Next, we visualized where each of the 12 categories appeared in the shared semantic space (Fig. 2a). Each category label was also assigned an RGB colour, where the red channel was determined by the first dimension, the green channel by the second, and the blue channel by the third. The first dimension is that which captured the most semantic variance across the voxel-wise models of all seven subjects. One end of this dimension favours categories related to humans and social interaction, including 'social', 'emotional', 'violent' and 'communal'. The other end favours categories related to perceptual descriptions, quantitative descriptions and setting, including 'tactile', 'locational', 'numeric' and 'visual'. This is consistent with previous suggestions that humans comprise a particularly salient and strongly represented semantic domain[16,21]. Subsequent dimensions of the semantic space captured less variance than the first and were also more difficult to interpret. The second dimension seems to distinguish between perceptual categories, including 'visual' and 'tactile', and non-perceptual categories, including 'mental', 'professional' and 'temporal'. The third and fourth dimensions are less clear.

Earlier studies identified the cortical regions comprising the semantic system[1,2], but could not comprehensively characterize their semantic

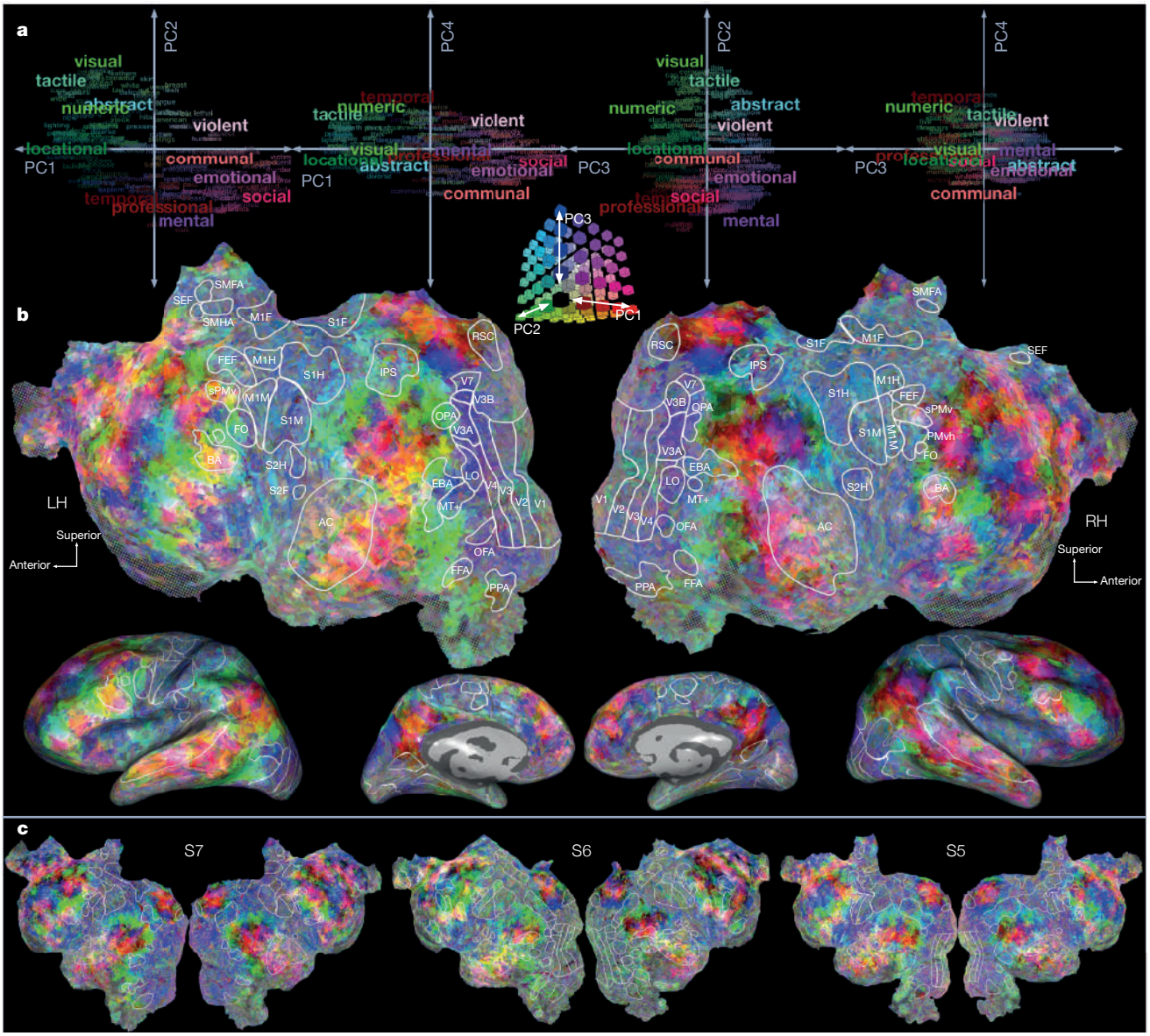


Figure 2 | Principal components of voxel-wise semantic models. **a**–**c**, Principal components analysis of voxel-wise model weights reveals four important semantic dimensions in the brain (Extended Data Fig. 2). **a**, An RGB colourmap was used to colour both words and voxels based on the first three dimensions of the semantic space. Words that best match the four semantic dimensions were found and then collapsed into 12 categories using *k*-means clustering. Each category (Supplementary Table 2) was manually assigned a label. The 12 category labels (large words) and a selection of the 458 best words (small words) are plotted here along four pairs of semantic dimensions. The largest axis of variation lies roughly along the first dimension, and separates perceptual and physical categories (tactile, locational) from human-related categories (social, emotional, violent). PC, principal component. **b**, Voxel-wise model weights were projected onto the semantic dimensions and then coloured using the same RGB colourmap (see Extended Data Fig. 3 for separate dimensions). Projections for one subject (S2) are shown on that subject's cortical surface. Semantic information seems to be represented in intricate patterns across much of the semantic system. **c**, Semantic principal component flatmaps for three other subjects. Comparing these flatmaps, many patterns appear to be shared across individuals. (See Extended Data Fig. 3 for other subjects.) Abbreviations for regions of interest are listed in the Methods section.

selectivity. We were able to visualize the pattern of semantic-domain selectivity across the entire cortex by projecting voxel-wise models onto the shared semantic dimensions. Figure 2b shows projections onto the first three dimensions for one subject, plotted together using the same RGB colour scheme as in Fig. 2a (Extended Data Fig. 3a shows each dimension separately). Thus, for example, a green voxel produces greater BOLD responses to categories that are coloured green in the semantic space, such as 'visual' and 'numeric'. This visualization suggests that semantic information is represented in intricate patterns that cover the semantic system, including broad regions of the prefrontal cortex, LTC and MTC, and LPC and MPC. Furthermore, these patterns appear to be relatively consistent across individuals (Fig. 2c; see also Extended Data Fig. 3b).

Using PrAGMATiC to construct a semantic atlas

Given the apparent consistency in the patterns of semantic selectivity across individuals, we sought to create a single atlas that describes the distribution of semantically selective functional areas in human cerebral cortex. To accomplish this, we developed a new Bayesian algorithm, PrAGMATiC, that produces a probabilistic and generative model of areas tiling the cortex[22]. This algorithm models patterns of functional tuning recovered by voxel-wise modelling as a dense, tiled map of functionally homogeneous brain areas (Fig. 3a), while respecting individual differences in anatomical and functional anatomy[23,24]. The arrangement and selectivity of these areas are determined by parameters learned from the fMRI data through a maximum-likelihood estimation technique similar to contrastive divergence[25].

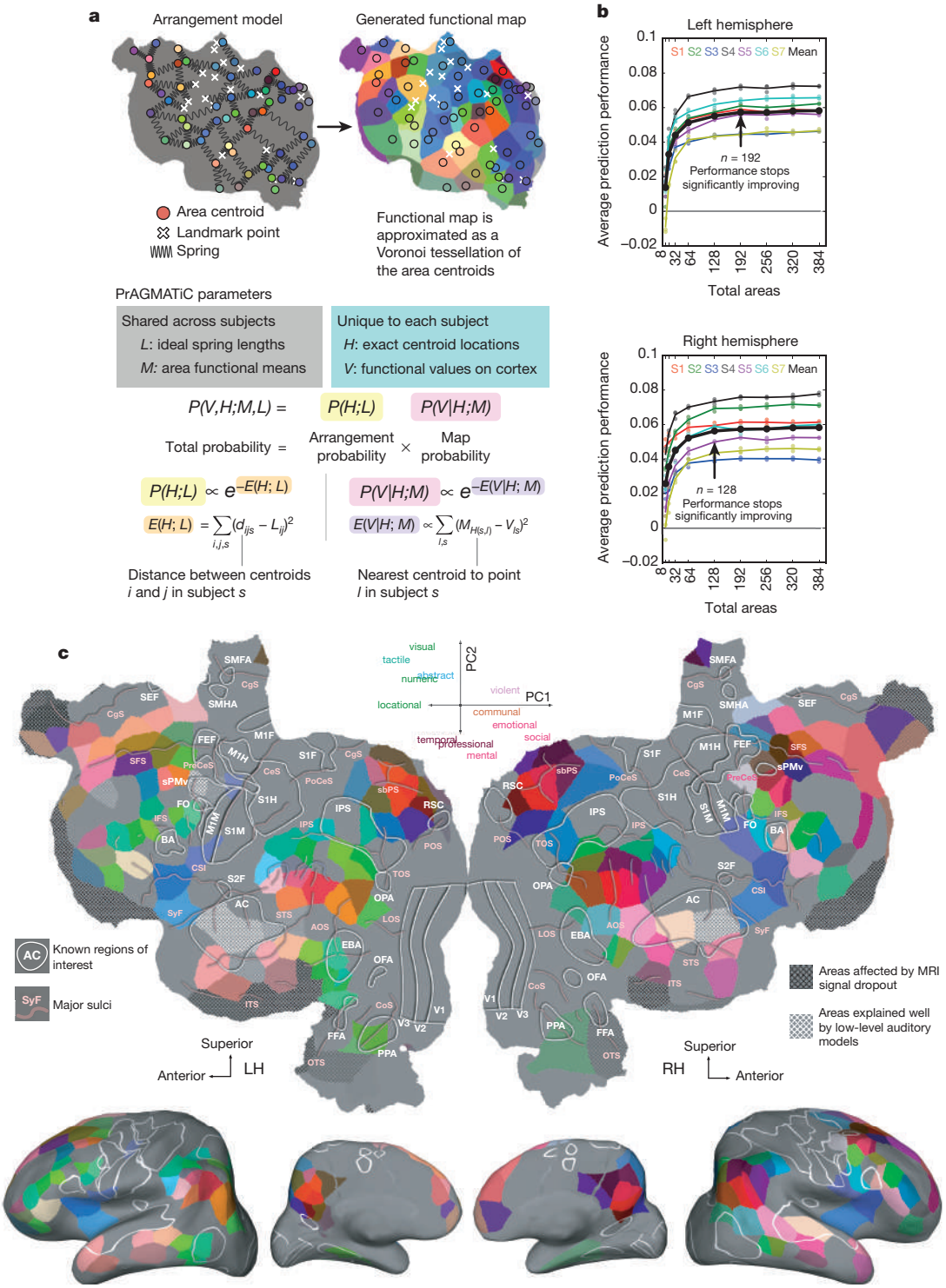


Figure 3 | PrAGMATiC: a generative model for cortical maps. **a**–**c**, To create an atlas that describes the distribution of semantically selective functional areas in the human cerebral cortex we developed PrAGMATiC, a probabilistic and generative model of areas tiling the cortex. **a**, PrAGMATiC has two parts: an arrangement model and an emission model. The arrangement model is analogous to a physical system of springs joining neighbouring area centroids. To enforce similarity across subjects, springs also join areas to 19 regions of interest that were localized separately. The emission model assigns the functional mean of the closest area centroid to each point on the cortex, forming a Voronoi tessellation. Spring lengths and area means are shared across subjects while exact area locations are unique to each subject. These parameters are fit using maximum-likelihood estimation. **b**, A leave-one-out procedure was used to choose the number of areas in each hemisphere. PrAGMATiC models were estimated on six subjects and then used to predict BOLD responses for the seventh. Prediction performance improved significantly up to 192 total areas in the left hemisphere and 128 areas in the right. **c**, A semantic atlas was estimated using data from all seven subjects. Areas for which the semantic model did not predict better than models based on low-level features (that is, word rate, phonemes) were removed. The remaining areas were plotted on one subject's cortical surface using the same RGB colourmap as Fig. 2. Areas dominated by signal dropout are shown in black hatching, and areas where the low-level models performed well are shown in white hatching. This atlas shows the functional organization of the semantic system that is common across subjects.

Some parameters are shared; these describe properties of the cortical map that are common across the group. Other parameters are unique to each subject; these capture individual differences. Learning both shared and unique parameters simultaneously eliminates the usual requirement to perform anatomical or functional alignment of data across subjects.

The PrAGMATiC algorithm has two components: an arrangement model that determines where functional areas appear on the cortical sheet, and an emission model that determines how the cortical map is produced from an arrangement of areas. The arrangement model simulates a physical spring network that joins the centroid of each functional area to its neighbours. Equilibrium spring lengths are shared across subjects, but each spring can be stretched or compressed in any individual subject. Arrangements are also constrained by several functional landmarks, which are known regions of interest identified in every subject using separate functional data. These constraints ensure that the maps will be similar across subjects, but allow for substantial individual variability in the precise arrangement and size of the areas. Using the arrangement model, the emission model creates homogeneous functional areas by assigning each vertex on the cortical surface to the nearest area centroid. The functional value at each vertex is then drawn from a multivariate normal distribution. The mean functional value for each area is learned by the algorithm and is shared across subjects. We define the functional value as a four-dimensional vector that reflects the projection of the estimated model for each voxel onto the four shared semantic dimensions.

One important hyperparameter is the total number of areas that PrAGMATiC uses to tile the cortex. We used a cross-validation procedure to choose the total number of areas tiling each hemisphere and then tested whether each area is semantically selective. PrAGMATiC models were estimated from data from six subjects and then used to predict the semantic map in the seventh subject using only cortical anatomy and the locations of functional landmarks in that subject. Predicted BOLD responses based on this map were compared to actual responses to determine how well the PrAGMATiC model generalizes across subjects. Prediction performance climbed quickly as the total number of areas rose from 8 to 128 and improved more gradually thereafter (Fig. 3b). In the left hemisphere, prediction performance did not improve significantly for models with 192 or more total areas (false discovery rate (FDR) > 0.01, Tukey post-hoc test with subject-wise random effects). In the right hemisphere, prediction performance did not improve significantly for models with 128 or more total areas. However, because PrAGMATiC tiles the entire cerebral cortex, these numbers include both semantically selective and nonselective areas. To identify the semantically selective areas and eliminate those that are nonselective, we tested whether the average voxel-wise semantic model in each area predicted responses significantly better than the average model for low-level features such as word rate, phoneme rate, and phonemes. This excluded areas that were not selective for either semantic or low-level features, such as motor and visual cortex. It also excluded areas that were not uniquely selective for semantic features, such as Broca's area, which was desirable because of the increased uncertainty of semantic model weights in those areas.

Figure 3c shows the semantic atlas projected onto the cortical surface of one subject (see also Extended Data Figs 4 and 5). The left hemisphere contains 77 semantic areas (FDR $< 1/192$, bootstrap test) and the right contains 63 semantic areas (FDR $< 1/128$, bootstrap test). A diverse tiling of areas that represent different semantic domains appear in the LPC (Extended Data Fig. 6), MPC (Extended Data Fig. 7) and SPFC (Extended Data Fig. 8). In the LPC and MPC, central areas (near the angular gyrus and subparietal sulci, respectively) are selective for social concepts, while surrounding areas are selective for numeric, visual or tactile concepts. In the SPFC, medial areas are mainly selective for social concepts, while dorsolateral areas are more diverse. The LPC, MPC and SPFC also all belong to the default mode network (DMN), which is thought to be involved in introspection, rumination and conscious thought[26]. One interesting possibility is that the semantic areas identified here represent the same semantic domains during conscious thought. This suggests that the contents of thought, or internal speech, might be decoded using these voxel-wise models[17]. In the LTC (Extended Data Fig. 9), our atlas identifies fewer distinct semantic areas than in the LPC, MPC or SPFC. This is surprising because the LTC has a key role in language comprehension[1,27] and also belongs to the DMN. However, the quality of fMRI signals recorded in the anterior temporal lobe is poor, so the LTC probably contains other semantic areas that could not be recovered using our current approach. Detailed analyses of semantic representations in the LPC, MPC, SPFC and LTC, as well as the VTC (Extended Data Fig. 10), IPFC (Extended Data Fig. 11), and opercular and insular cortex (Extended Data Fig. 12) can be found in Supplementary Information, along with discussion and comparisons to earlier neuroimaging and lesion results.

Discussion

One striking aspect of our atlas is that the distribution of semantically selective areas is relatively symmetrical across the two cerebral hemispheres. This finding is inconsistent with human lesion studies that support the idea that semantic representation is lateralized to the left hemisphere[13]. However, many fMRI studies of semantic representation find only modest lateralization[1] and one study that used narrative stories found highly bilateral results similar to ours[2]. This suggests that right hemisphere areas may respond more strongly to narrative stimuli than to the words and short phrases used in most studies. Still, more research will be needed to determine what roles these left- and right-hemisphere semantic areas have in language comprehension.

Another interesting aspect of these results is that the organization of semantically selective brain areas seems to be highly consistent across individuals. This might suggest that innate anatomical connectivity or cortical cytoarchitecture constrains the organization of high-level semantic representations[28,29]. It is also possible that this is owing to common life experiences of the subjects, all of whom were raised and educated in Western industrial societies. Future studies that include subjects from more diverse backgrounds will be needed to determine how much of this organizational consistency reflects innate brain structure versus experience.

One limitation of PrAGMATiC as used here is that each area is assumed to be functionally homogeneous. This is a common assumption in the design and analysis of many neuroimaging studies[30]. However, many cortical maps, including semantic maps in visual cortex[14], seem to contain smoothly changing gradients of representation. It should be possible to modify the PrAGMATiC algorithm to model functional gradients explicitly. This will provide an objective tool for determining whether the semantic maps found here are best described as homogeneous areas or as gradients.

Data-driven approaches are commonplace in studies of human neuroanatomy[31] and resting state networks[26,32], but are only beginning to be used in functional imaging[14,15]. Our study demonstrates the power and efficiency of data-driven approaches for functional mapping of the human brain. Although our experiment used a simple design in which subjects only listened to stories, the data were rich enough to produce a comprehensive atlas of semantically selective areas. Furthermore, our data-driven framework is quite general. Other properties of language can be mapped (even in this same data set) by using feature spaces that reflect phonemes, syntax and so on. Complex semantic models that incorporate information beyond word co-occurrence can be tested and compared quantitatively. The generalizability of these models can also be tested by using stimuli beyond autobiographical stories. It is sometimes difficult to synthesize the results of data-driven experiments with those from hypothesis-driven experiments, but future methodological and theoretical developments should help to bridge this divide. We expect that the semantic atlas presented here will be useful for many researchers investigating the neurobiological basis of language. We also expect that this atlas can be refined and expanded

by incorporating results from future studies. To facilitate this, we have created a detailed interactive version of the semantic atlas that can be explored online at http://gallantlab.org/huth2016.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 January 2014; accepted 2 March 2016.

1. Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* **19,** 2767–2796 (2009).
2. Lerner, Y., Honey, C. J., Silbert, L. J. & Hasson, U. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J. Neurosci.* **31,** 2906–2915 (2011).
3. Friederici, A. D., Opitz, B. & von Cramon, D. Y. Segregating semantic and syntactic aspects of processing in the human brain: an fMRI investigation of different word types. *Cereb. Cortex* **10,** 698–705 (2000).
4. Noppeney, U. & Price, C. J. Retrieval of abstract semantics. *Neuroimage* **22,** 164–170 (2004).
5. Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T. & Medler, D. A. Distinct brain systems for processing concrete and abstract concepts. *J. Cogn. Neurosci.* **17,** 905–917 (2005).
6. Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A. & Saxe, R. Concepts are more than percepts: the case of action verbs. *J. Neurosci.* **28,** 11347–11353 (2008).
7. Saxe, R. & Kanwisher, N. People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage* **19,** 1835–1842 (2003).
8. Caramazza, A. & Shelton, J. R. Domain-specific knowledge systems in the brain the animate-inanimate distinction. *J. Cogn. Neurosci.* **10,** 1–34 (1998).
9. Mummery, C. J., Patterson, K., Hodges, J. R. & Price, C. J. Functional neuroanatomy of the semantic system: divisible by what? *J. Cogn. Neurosci.* **10,** 766–777 (1998).
10. Just, M. A., Cherkassky, V. L., Aryal, S. & Mitchell, T. M. A neurosemantic theory of concrete noun representation based on the underlying brain codes. *PLoS ONE* **5,** e8622 (2010).
11. Warrington, E. K. The selective impairment of semantic memory. *Q. J. Exp. Psychol.* **27,** 635–657 (1975).
12. Mitchell, T. M. *et al.* Predicting human brain activity associated with the meanings of nouns. *Science* **320,** 1191–1195 (2008).
13. Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D. & Damasio, A. R. A neural basis for lexical retrieval. *Nature* **380,** 499–505 (1996).
14. Huth, A. G., Nishimoto, S., Vu, A. T. & Gallant, J. L. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron* **76,** 1210–1224 (2012).
15. Wehbe, L. *et al.* Simultaneously uncovering the patterns of brain regions involved in different story reading subprocesses. *PLoS ONE* **9,** e112575 (2014).
16. Naselaris, T., Prenger, R. J., Kay, K. N., Oliver, M. & Gallant, J. L. Bayesian reconstruction of natural images from human brain activity. *Neuron* **63,** 902–915 (2009).
17. Nishimoto, S. *et al.* Reconstructing visual experiences from brain activity evoked by natural movies. *Curr. Biol.* **21,** 1641–1646 (2011).
18. Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. & Harshman, R. Indexing by latent semantic analysis. *J. Am. Soc. Inf. Sci.* **41,** 391–407 (1990).
19. Lund, K. & Burgess, C. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behav. Res. Methods Instrum. Comput.* **28,** 203–208 (1996).
20. Turney, P. D. & Pantel, P. From frequency to meaning: vector space models of semantics. *J. Artif. Intell. Res.* **37,** 141–188 (2010).
21. Caramazza, A. & Mahon, B. Z. The organisation of conceptual knowledge in the brain: the future's past and some future directions. *Cogn. Neuropsychol.* **23,** 13–38 (2006).
22. Huth, A. G., Griffiths, T. L., Theunissen, F. E. & Gallant, J. L. PrAGMATiC: a probabilistic and generative model of areas tiling the cortex. Preprint at http://arxiv.org/abs/1504.03622 (2015).
23. Amunts, K., Malikovic, A., Mohlberg, H., Schormann, T. & Zilles, K. Brodmann's areas 17 and 18 brought into stereotaxic space—where and how variable? *Neuroimage* **11,** 66–84 (2000).
24. Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S. & Kanwisher, N. New method for fMRI investigations of language: defining ROIs functionally in individual subjects. *J. Neurophysiol.* **104,** 1177–1194 (2010).
25. Hinton, G. E. Training products of experts by minimizing contrastive divergence. *Neural Comput.* **14,** 1771–1800 (2002).
26. Buckner, R. L., Andrews-Hanna, J. R. & Schacter, D. L. The brain's default network: anatomy, function, and relevance to disease. *Ann. NY Acad. Sci.* **1124,** 1–38 (2008).
27. DeWitt, I. & Rauschecker, J. P. Phoneme and word recognition in the auditory ventral stream. *Proc. Natl Acad. Sci. USA* **109,** E505–E514 (2012).
28. Riesenhuber, M. Appearance isn't everything: news on object representation in cortex. *Neuron* **55,** 341–344 (2007).
29. Dehaene, S., Cohen, L., Sigman, M. & Vinckier, F. The neural code for written words: a proposal. *Trends Cogn. Sci.* **9,** 335–341 (2005).
30. Op de Beeck, H. P., Haushofer, J. & Kanwisher, N. G. Interpreting fMRI data: maps, modules and dimensions. *Nature Rev. Neurosci.* **9,** 123–135 (2008).
31. Caspers, S. *et al.* Organization of the human inferior parietal lobule based on receptor architectonics. *Cereb. Cortex* **23,** 615–628 (2013).
32. Cohen, A. L. *et al.* Defining functional areas in individual human brains using resting functional connectivity MRI. *Neuroimage* **41,** 45–57 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by grants from the National Science Foundation (NSF; IIS1208203), the National Eye Institute (EY019684), and from the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370. A.G.H. was also supported by the William Orr Dingwall Neurolinguistics Fellowship. We thank J. Sohl-Dickstein and K. Crane for technical discussions about PrAGMATiC, J. Nguyen for assistance transcribing and aligning stimuli, B. Griffin for segmenting and flattening cortical surfaces, and N. Bilenko, J. Gao, M. Lescroart and A. Nunez-Elizalde for general comments and discussions.

Author Contributions All authors helped conceive and design the experiment. W.A.d.H. and A.G.H. selected and annotated stimuli and collected fMRI data. A.G.H. analysed the data. A.G.H. and T.L.G. designed the PrAGMATiC generative model. A.G.H. and J.L.G. wrote the paper. J.L.G. contributed to all aspects of the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.L.G. (gallant@berkeley.edu).

METHODS

MRI data collection. MRI data were collected on a 3T Siemens TIM Trio scanner at the UC Berkeley Brain Imaging Center using a 32-channel Siemens volume coil. Functional scans were collected using gradient echo EPI with repetition time (TR) = 2.0045 s, echo time (TE) = 31 ms, flip angle = 70°, voxel size = 2.24 × 2.24 × 4.1 mm (slice thickness = 3.5 mm with 18% slice gap), matrix size = 100 × 100, and field of view = 224 × 224 mm. Thirty axial slices were prescribed to cover the entire cortex and were scanned in interleaved order. A custom-modified bipolar water excitation radiofrequency (RF) pulse was used to avoid signal from fat. Anatomical data were collected using a T1-weighted multi-echo MP-RAGE sequence on the same 3T scanner.

Subjects. Functional data were collected from five male subjects and two female subjects: S1 (male, age 26), S2 (male, age 32), S3 (female, age 31), S4 (male, age 31), S5 (male, age 26), S6 (female, age 25), and S7 (male, age 30). Two of the subjects were authors (S1: A.G.H.; and S3: W.A.d.H.). All subjects were healthy and had normal hearing. The experimental protocol was approved by the Committee for the Protection of Human Subjects at University of California, Berkeley. Written informed consent was obtained from all subjects. Voxel-wise models were estimated and validated independently for each subject using separate data sets reserved for that purpose. Principal components and PrAGMATiC analyses used leave-one-subject-out cross-validation to verify that the group models accurately predict the data recorded in each individual subject.

Natural story stimuli. The model estimation data set consisted of ten 10- to 15-min stories taken from *The Moth Radio Hour*. In each story, a single speaker tells an autobiographical story in front of a live audience. The ten selected stories cover a wide range of topics and are highly engaging. Each story was played during a separate fMRI scan. The length of each scan was tailored to the story, and included 10 s of silence both before and after the story. These data were collected during two 2-h scanning sessions that were performed on different days. The model validation data set consisted of one 10-min story, also taken from *The Moth Radio Hour*. This story was played twice for each subject (once during each scanning session), and then the two responses were averaged. For story synopses and details of story transcription and preprocessing procedures, see Supplementary Methods.

Stories were played over Sensimetrics S14 in-ear piezoelectric headphones. A Behringer Ultra-Curve Pro hardware parametric equalizer was used to flatten the frequency response of the headphones based on calibration data provided by Sensimetrics. All stimuli were played at 44.1 kHz using the pygame library in Python. All stimuli were normalized to have peak loudness of −1 dB relative to maximum. However, the stories were performed by different speakers and were not uniformly mastered, so some differences in total loudness remain.

Story transcription and preprocessing. Each story was manually transcribed by one listener, and then the transcript was checked by a second listener. Certain sounds (for example, laughter, lip-smacking and breathing) were also marked to improve the accuracy of the automated alignment. The audio of each story was downsampled to 11 kHz and the Penn Phonetics Lab Forced Aligner (P2FA[33]) was used to automatically align the audio to the transcript. The forced aligner uses a phonetic hidden Markov model to find the temporal onset and offset of each word and phoneme. The Carnegie Mellon University (CMU) pronouncing dictionary was used to guess the pronunciation of each word. When necessary, words and word fragments that appeared in the transcript but not in the dictionary were manually added. After automatic alignment was complete, Praat[34] was used to check and correct each aligned transcript manually. The corrected aligned transcript was then spot-checked for accuracy by a different listener.

Finally, the aligned transcripts were converted into separate word and phoneme representations. The phoneme representation of each story is a list of pairs (p,t), where p is a phoneme and t is the time from the beginning of the story to the middle of the phoneme (that is, halfway between the start and end of the phoneme) in seconds. Similarly the word representation of each story is a list of pairs (w,t), where w is a word.

Semantic model construction. To account for response variance caused by the semantic content of the stories, we constructed a 985-dimensional semantic feature space based on word co-occurrence statistics in a large corpus of text[12,18,19]. First, we constructed a 10,470-word lexicon from the union of the set of all words appearing in the stories and the 10,000 most common words in the large text corpus. We then selected 985 basis words from Wikipedia's *List of 1000 Basic Words* (contrary to the title, this list contained only 985 unique words at the time it was accessed). This basis set was selected because it consists of common words that span a very broad range of topics. The text corpus used to construct this feature space includes the transcripts of 13 *Moth* stories (including the 10 used as stimuli in this experiment), 604 popular books, 2,405,569 Wikipedia pages, and 36,333,459 user comments scraped from reddit.com. In total, the 10,470 words in our lexicon appeared 1,548,774,960 times in this corpus.

Next, we constructed a word co-occurrence matrix, M, with 985 rows and 10,470 columns. Iterating through the text corpus, we added 1 to $M_{i,j}$ each time word j appeared within 15 words of basis word i. A window size of 15 was selected to be large enough to suppress syntactic effects (that is, word order) but no larger. Once the word co-occurrence matrix was complete, we log-transformed the counts, replacing $M_{i,j}$ with $\log(1 + M_{i,j})$. Next, each row of M was z-scored to correct for differences in basis word frequency, and then each column of M was z-scored to correct for word frequency. Each column of M is now a 985-dimensional semantic vector representing one word in the lexicon.

The matrix used for voxel-wise model estimation was then constructed from the stories: for each word–time pair (w,t) in each story we selected the corresponding column of M, creating a new list of semantic vector–time pairs, $(\mathbf{M}_w,t)$. These vectors were then resampled at times corresponding to the fMRI acquisitions using a 3-lobe Lanczos filter with the cut-off frequency set to the Nyquist frequency of the fMRI acquisition (0.249 Hz).

Voxel-wise model estimation and validation. A linearized finite impulse response (FIR) model[14,17] consisting of four separate feature spaces was fit to every cortical voxel in each subject's brain. These four feature spaces were word rate (1 feature), phoneme rate (1 feature), phonemes (39 features), and semantics (985 features). The word rate, phoneme rate, and phoneme features were used to account for responses to low-level properties of the stories that could contaminate the semantic model weights (see Supplementary Methods for details of how these low-level models were constructed). A separate linear temporal filter with four delays (1, 2, 3, and 4 time points) was fit for each of these 1,026 features, yielding a total of 4,104 features. This was accomplished by concatenating feature vectors that had been delayed by 1, 2, 3, and 4 time points (2, 4, 6, and 8 s). Thus, in the concatenated feature space one channel represents the word rate 2 s earlier, another 4 s earlier, and so on. Taking the dot product of this concatenated feature space with a set of linear weights is functionally equivalent to convolving the original stimulus vectors with linear temporal kernels that have non-zero entries for 1-, 2-, 3-, and 4-time-point delays.

Before doing regression, we first z-scored each feature channel within each story. This was done to match the features to the fMRI responses, which were also z-scored within each story. However, this had little effect on the learned weights.

The 4,104 weights for each voxel were estimated using L2-regularized linear regression (also known as ridge regression). To keep the scale of the weights consistent and to prevent bias in subsequent analyses, a single value of the regularization coefficient was used for all voxels in all subjects. This regularization coefficient was found by bootstrapping the regression procedure 50 times in each subject. In each bootstrap iteration, 800 time points (20 blocks of 40 consecutive time points each) were removed from the model estimation data set and reserved for testing. Then the model weights were estimated on the remaining 2,937 time points for each of 20 possible regularization coefficients (log spaced between 10 and 1,000). These weights were used to predict responses for the 800 reserved time points, and then the correlation between actual and predicted responses was found. After the bootstrapping was complete, a regularization–performance curve was obtained for each subject by averaging the bootstrap sample correlations first across the 50 samples and then across all voxels. Next, the regularization–performance curves were averaged across the seven subjects and the best overall value of the regularization parameter (183.3) was selected. The best overall regularization parameter value was also the best value in three individual subjects. For the other four subjects the best regularization parameter value was slightly higher (233.6).

To validate the voxel-wise models, estimated semantic feature weights were used to predict responses to a separate story that had not been used for weight estimation. Prediction performance was then estimated as the Pearson correlation between predicted and actual responses for each voxel over the 290 time points in the validation story. Statistical significance was computed by comparing estimated correlations to the null distribution of correlations between two independent Gaussian random vectors of the same length. Resulting P values were corrected for multiple comparisons within each subject using the false discovery rate (FDR) procedure[35].

All model fitting and analysis was performed using custom software written in Python, making heavy use of NumPy[36], SciPy[37], and pycortex[38].

Semantic principal components analysis. We used principal components analysis (PCA) to recover a low-dimensional semantic space from the estimated semantic model weights. We first selected only the 10,000 best predicted voxels in each subject according to the average bootstrap correlation (for the selected regularization parameter value) obtained during model estimation. This was done to avoid including noise from poorly modelled voxels. Then we removed temporal information from the voxel-wise model weights by averaging across the four delays for each feature. The weights for the word frequency, phoneme frequency, and phoneme features were then discarded, leaving only the 985 semantic model weights for each voxel. Finally, we applied PCA to these weights, yielding 985 principal components

(PCs). Partial scree plots showing the amount of variance accounted for by each PC are shown in Extended Data Fig. 2. See Supplementary Methods for details.

PrAGMATiC. The PrAGMATiC generative model[22] has two components: an arrangement model and an emission model. The arrangement model defines a probability distribution over possible arrangements of the functional areas. This model assumes that the location of each area is defined by a single point called the area centroid. Each centroid is modelled as being joined to nearby centroids by springs. While exact centroid locations can vary from subject to subject, the equilibrium length of each spring is assumed to be consistent across subjects. The probability distribution over possible locations of the centroids is defined using the total potential energy of the spring system. This distribution assigns a high probability to low-energy arrangements of the centroids (that is, where the springs are not stretched much and so store little potential energy) and low probability to high-energy arrangements (where the springs are stretched a lot).

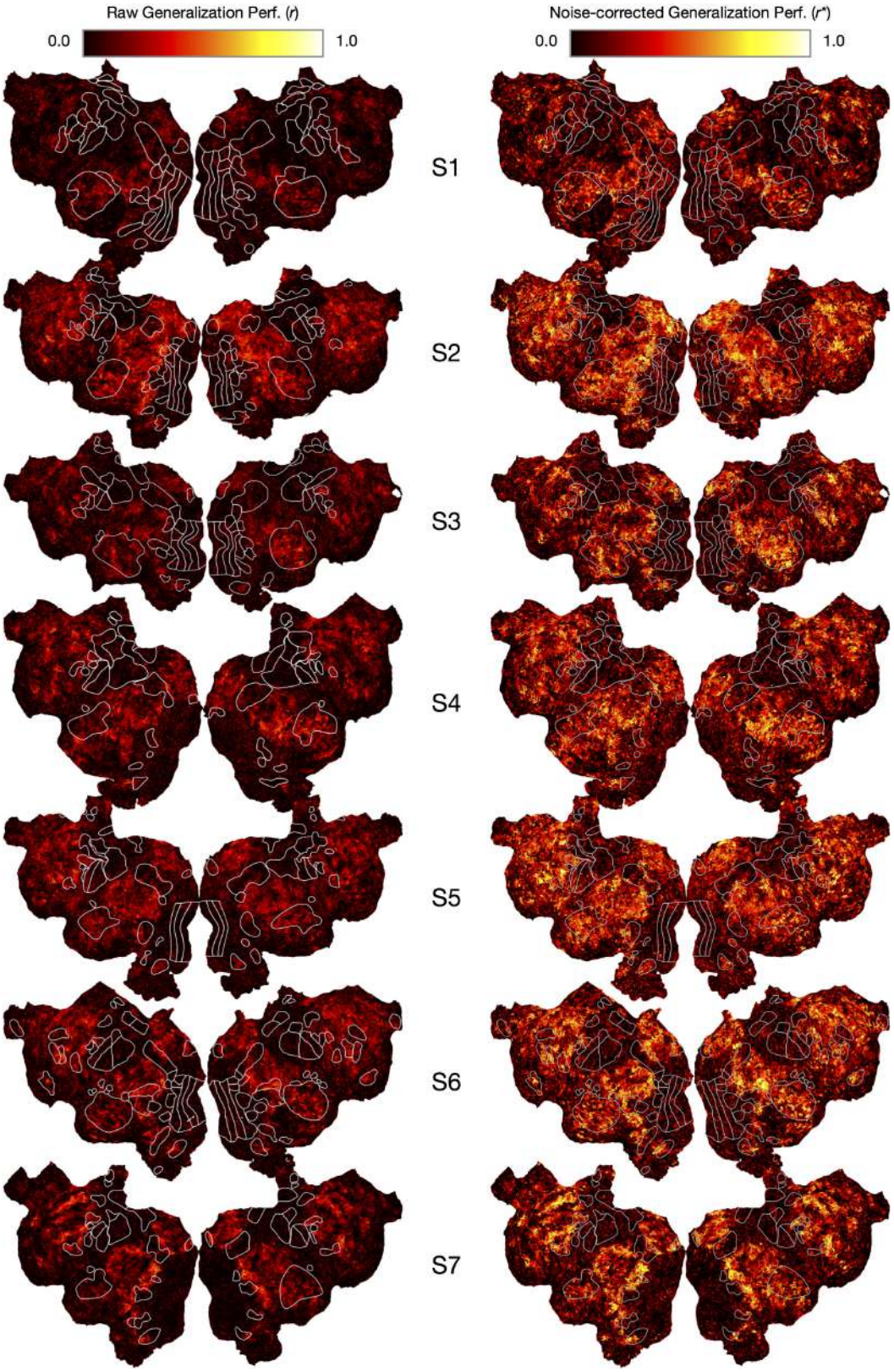
The second component is the emission model, which defines a probability distribution over semantic maps given an arrangement of functional areas. In the emission model each area centroid is assigned a particular semantic value in the four-dimensional common semantic space. This value determines what type of semantic information is represented in that area. To generate a semantic map from any particular arrangement, each point on the cortical surface is first assigned to the closest area centroid (creating a Voronoi diagram). Then the semantic value for each point is sampled from a spherical Gaussian distribution in semantic space, centred on the semantic value of the centroid.

A consequence of modelling semantic maps using a Voronoi diagram is that every point on the cortex must be assigned to an area, while we know that many points on the cortex are not semantically selective. We distinguished between semantically selective and non-selective areas by testing whether the mean semantic voxel-wise model in each area predicted responses significantly better on a held-out story than a baseline model that accounts for responses to phonemes and word rate.

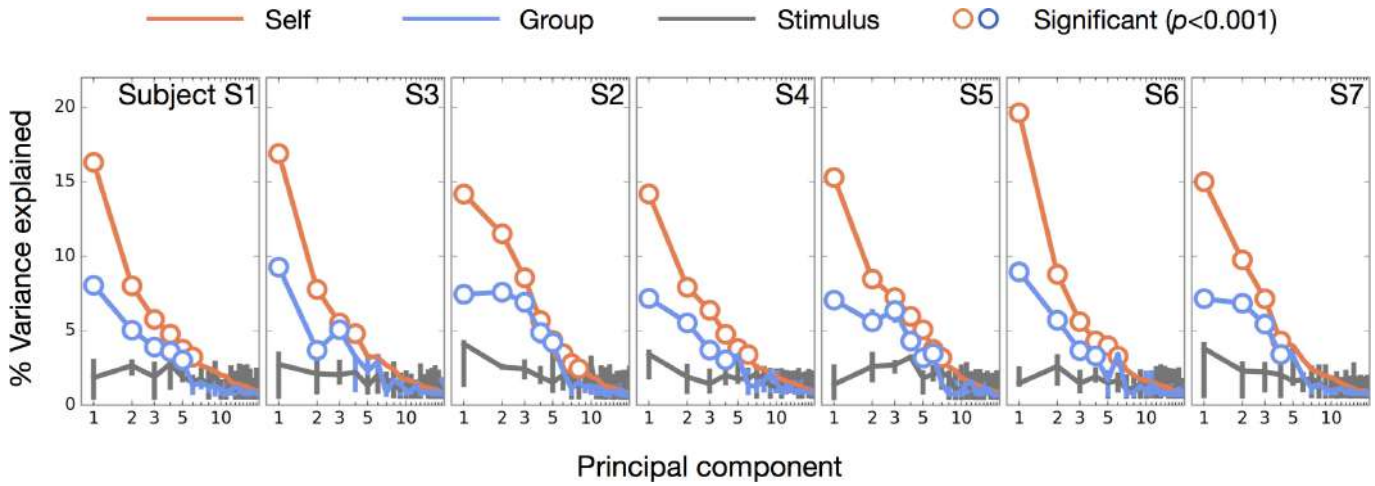
To train the generative model we derived maximum-likelihood estimation (MLE) update rules similar to the Boltzmann learning rule with contrastive divergence[25]. We used these learning rules to iteratively update the spring lengths and semantic values, maximizing the probability of the observed maps and minimizing the probability of unobserved maps. For details see Supplementary Methods.

Region of interest abbreviations. Fusiform face area (FFA), occipital face area (OFA), parahippocampal place area (PPA), occipital place area (OPA), retrosplenial cortex (RSC), extrastriate body area (EBA), visual areas (V1-V4, V3A, V3B, V7), lateral occipital visual area (LO), middle temporal visual area (MT+), intraparietal sulcus visual area (IPS), auditory cortex (AC), primary motor and somatosensory areas for feet (M1F, S1F), hands (M1H, S1H), and mouth (M1M, S1M), secondary somatosensory areas for feet (S2F), and hands (S2H), frontal eye fields (FEF), frontal opercular eye movement area (FO), supplementary motor foot area (SMFA), and hand area (SMHA), supplementary eye fields (SEF), Broca's area (BA), superior premotor ventral speech area (sPMv), premotor ventral hand area (PMvh).

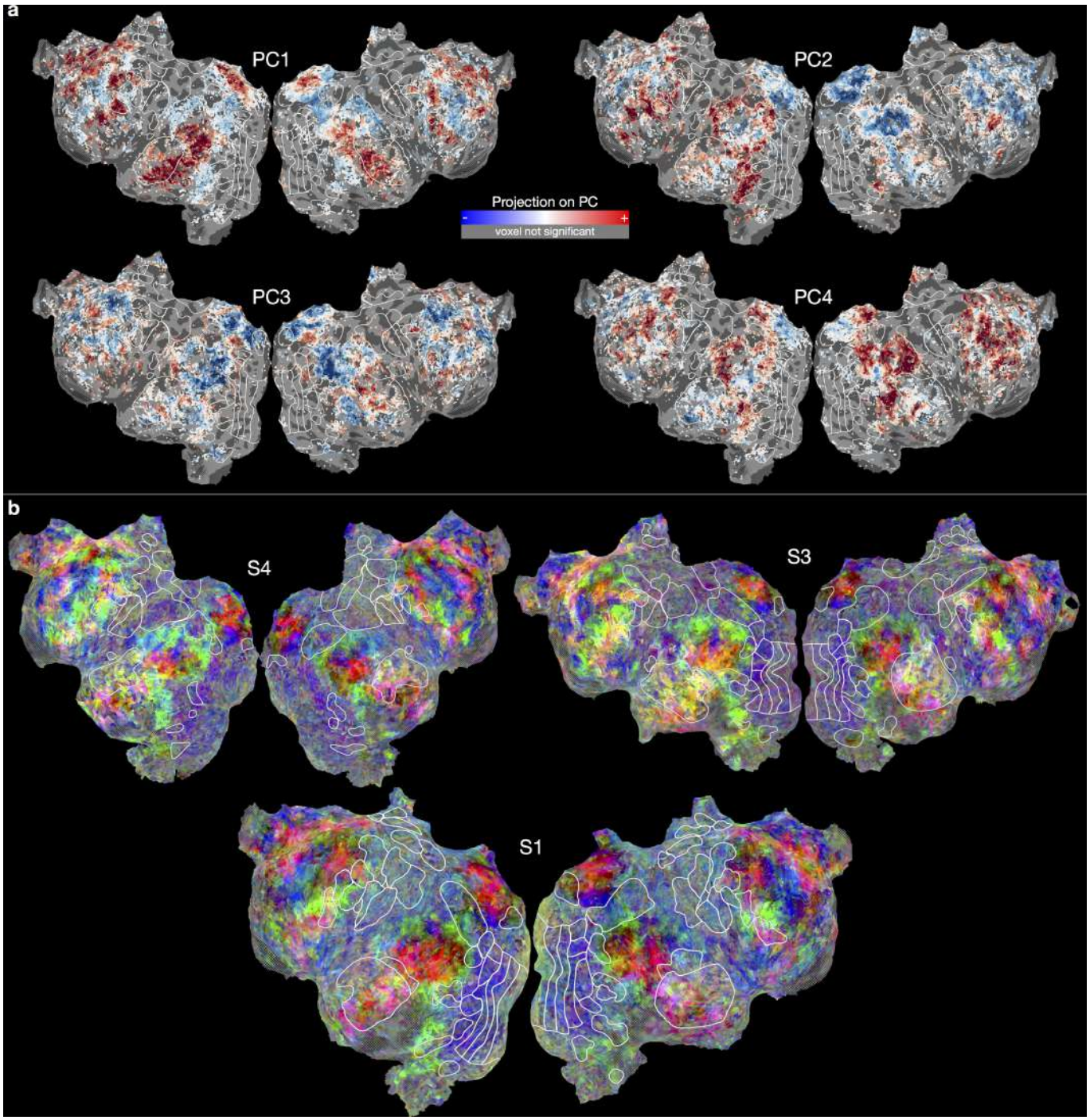
33. Yuan, J. & Liberman, M. Speaker identification on the SCOTUS corpus. *Proc. Acoust.* Preprint at http://www.ling.upenn.edu/~jiahong/publications/c09.pdf (2008).
34. Boersma, P. & Weenink, D. Praat: doing phonetics by computer (University of Amsterdam, 2014).
35. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* **57,** 289–300 (1995).
36. Oliphant, T. E. *Guide to NumPy* (Brigham Young University, 2006).
37. Jones, E., Oliphant, T. E. & Peterson, P. SciPy: Open source scientific tools for Python (SciPy, 2001).
38. Gao, J. S., Huth, A. G., Lescroart, M. D. & Gallant, J. L. Pycortex: an interactive surface visualizer for fMRI. *Front. Neuroinform.* **9,** 23 (2015).



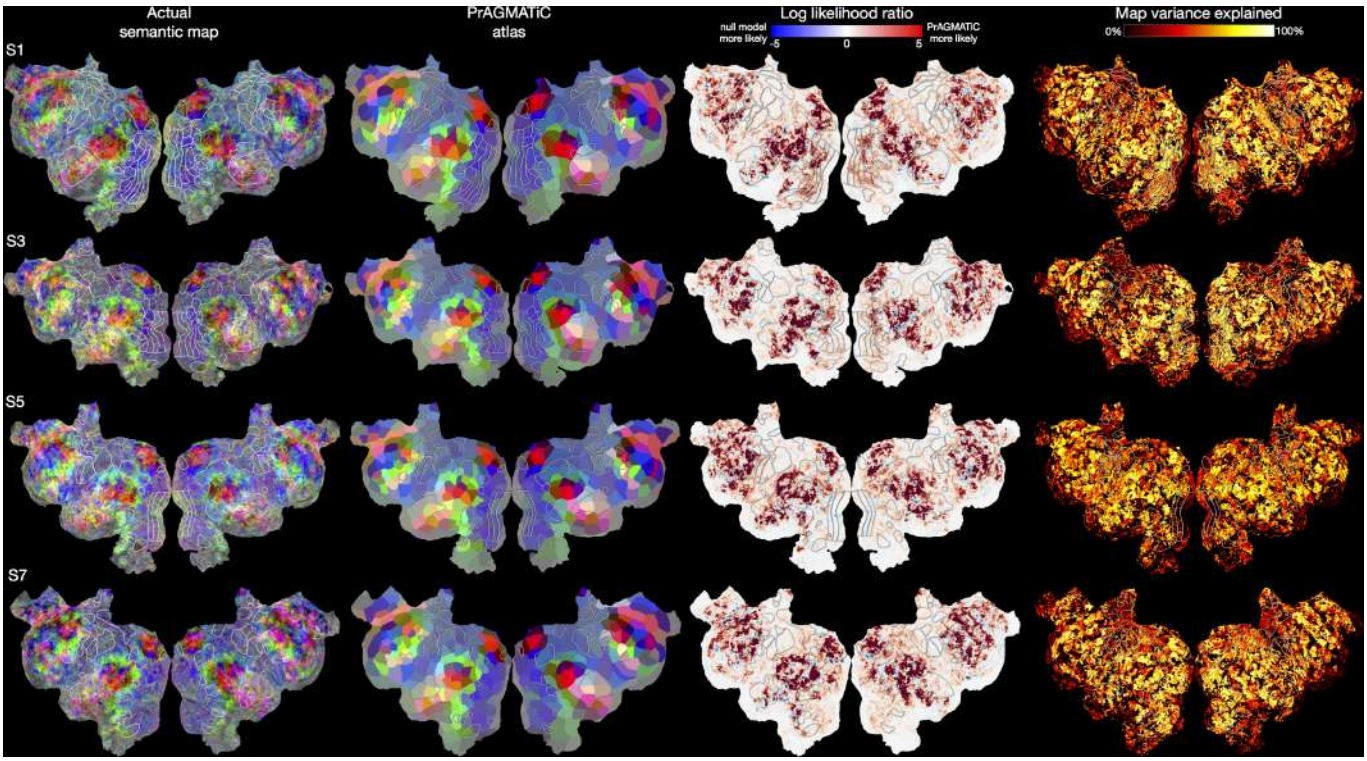
Extended Data Figure 1 | Voxel-wise model prediction performance. Cortical flatmaps showing prediction performance of voxel-wise semantic models for all seven subjects, formatted similarly to Fig. 1c. Models were tested using one 10-min story that was not included during model estimation. Prediction performance was then computed as the correlation between predicted and measured BOLD responses. Left column, raw prediction performance. Note that the colourmap here is scaled 0–1 rather than 0–0.6 as in Fig. 1c to match the scale of the adjusted prediction performance maps. Right column, prediction performance corrected to account for different amounts of noise in the BOLD responses (see Supplementary Methods for details). The voxel-wise semantic models predict BOLD responses in many brain areas, including SPFC, IPFC, LTC, VTC, LPC and MPC. These same regions have been previously identified as the semantic system in the human brain.



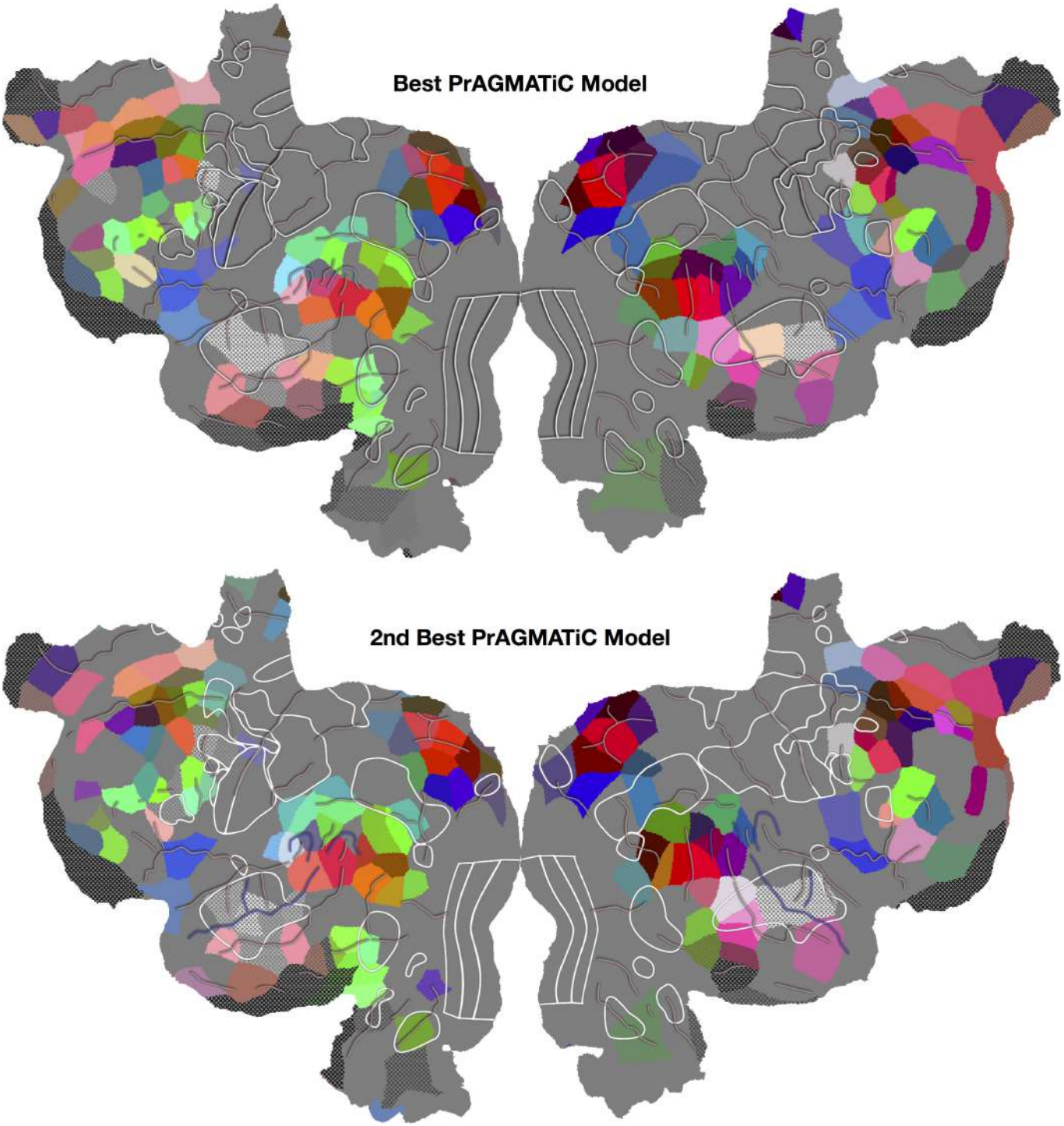
Extended Data Figure 2 | Amount of variance explained by individual subject and group semantic dimensions. Principal components analysis was used to discover the most important semantic dimensions from voxel-wise semantic model weights in each subject. To reduce noise, we used only the 10,000 best voxels in each subject, determined by cross-validation within the model estimation data set. Here we show the amount of variance explained in the semantic model weights by each of the 20 most important principal components (PCs). Orange lines show the amount of variance explained by each subject's own PCs, blue lines show the variance explained by the PCs of combined data from the other six subjects, and grey lines show the variance explained by the PCs of the stories. (The Gale–Shapley stable marriage algorithm was used to re-order the group and stimulus PCs to maximize their correlation with the subject's PCs.) Error bars indicate 99% confidence intervals. Confidence intervals for the subjects' own PCs and group PCs are very small. Hollow markers indicate subject or group PCs that explain significantly more variance than the corresponding stimulus PCs ($P < 0.001$, bootstrap test). Six PCs explain significantly more variance in one out of seven subjects, five PCs in two subjects, four PCs in three subjects, and three PCs in one subject. Thus, four PCs seem to comprise a semantic space that is common across most individuals.



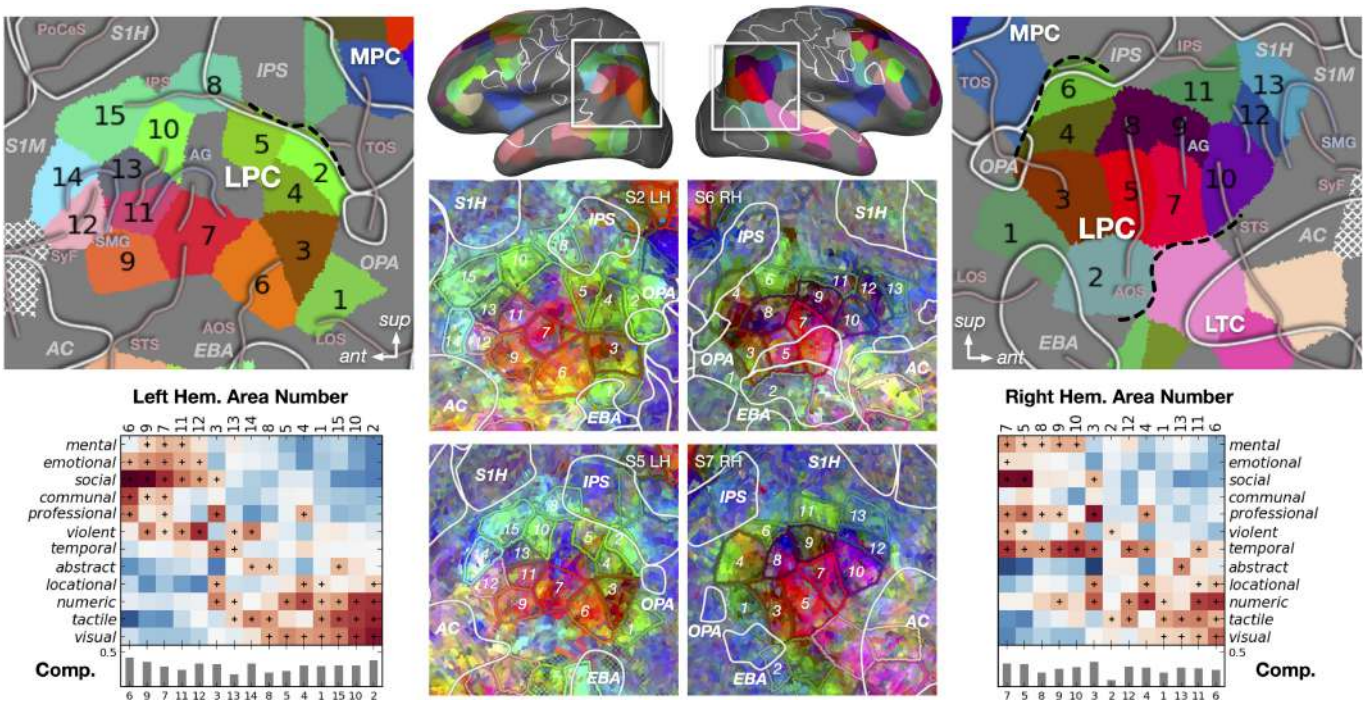
Extended Data Figure 3 | Separate cortical projections of semantic dimensions 1–4 on subject S2 and combined cortical projections of dimensions 1–3 for subjects S1, S3 and S4. a, Voxel-wise semantic model weights for subject S2 were projected onto each of the common semantic dimensions defined by PCs 1–4. Voxels for which model generalization performance was not significantly greater than zero (q(FDR) > 0.05) are shown in grey. Positive projections are shown in red, negative projections in blue and near-zero projections in white. Voxels with fMRI signal dropout due to field inhomogeneity are shaded with black hatched lines. **b**. Like Fig. 2b, c, this panel shows the result of projecting voxel-wise models onto the first three common semantic dimensions, and then colouring each voxel using an RGB colourmap. The red colour component corresponds to the projection on the first PC, the green component to the second, and the blue component to the third. Semantic information seems to be represented in complex patterns distributed across the semantic system and the patterns seem to be largely conserved across individuals.



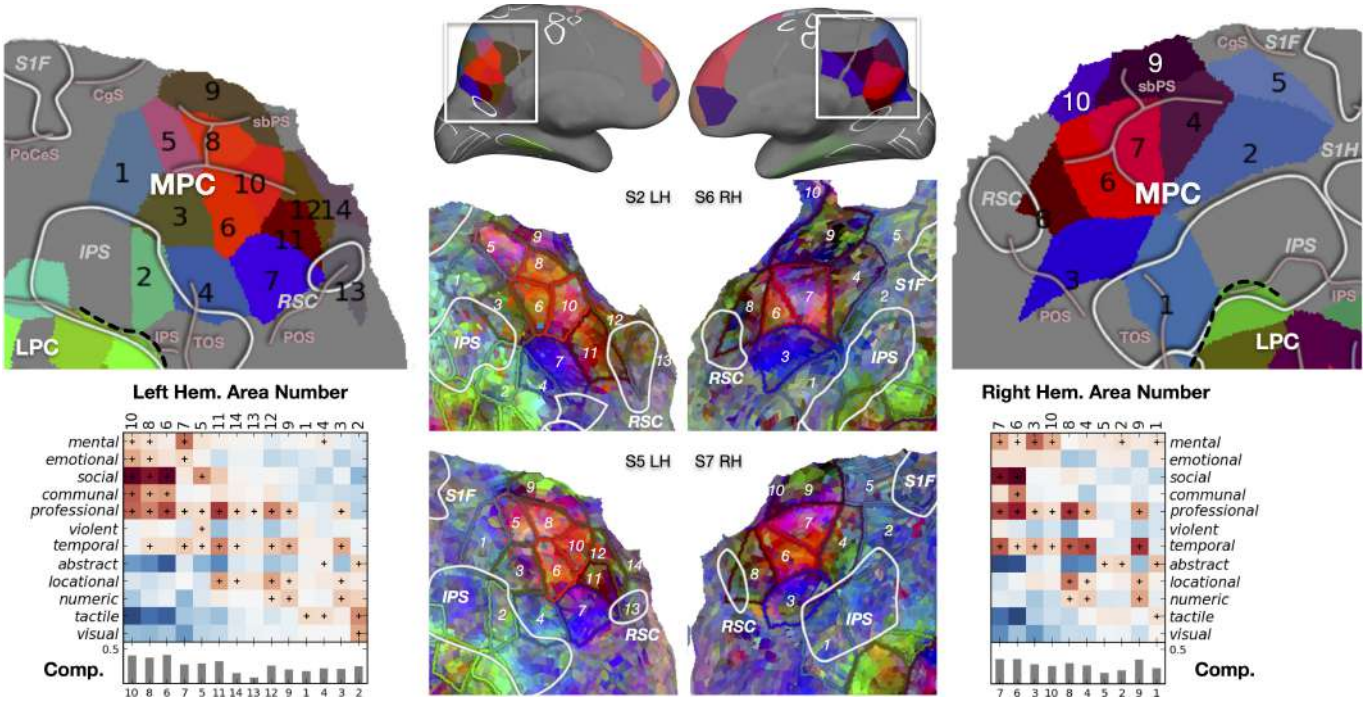
Extended Data Figure 4 | PrAGMATiC atlas likelihood maps. Comparison of actual semantic maps (Fig. 2, Extended Data Fig. 3) to the maps generated from the PrAGMATiC atlas (Fig. 3). PrAGMATiC atlases for the left and right hemispheres were fit using data from all seven subjects. The left hemisphere atlas has 192 total areas and the right hemisphere has 128 (including non-semantic areas). Here we show the actual semantic maps for four subjects (first column), the PrAGMATiC atlas on each subject's cortical surface (second column), the log likelihood ratio of the actual semantic map under the PrAGMATiC atlas versus a null model (third column), and the fraction of variance in the semantic map that the PrAGMATiC atlas explains for each location on the cortical surface (fourth column). The likelihood ratio maps show that most areas where there are large semantic model weights (that is, the semantic system) are much better explained by PrAGMATiC than by a null model and thus appear red, while areas where the weights are small (that is, somatomotor cortex, visual cortex, and so on) are about equally well explained by both PrAGMATiC and the null model and thus appear white. Variance explained was computed by subtracting the PrAGMATiC atlas from the actual semantic map (in the space of the four group semantic dimensions), squaring and summing the residuals and then dividing by the sum of squares in the actual map. The variance explained maps show that the PrAGMATiC atlas captures a large fraction of the variance in the semantic maps (37–47% in total).



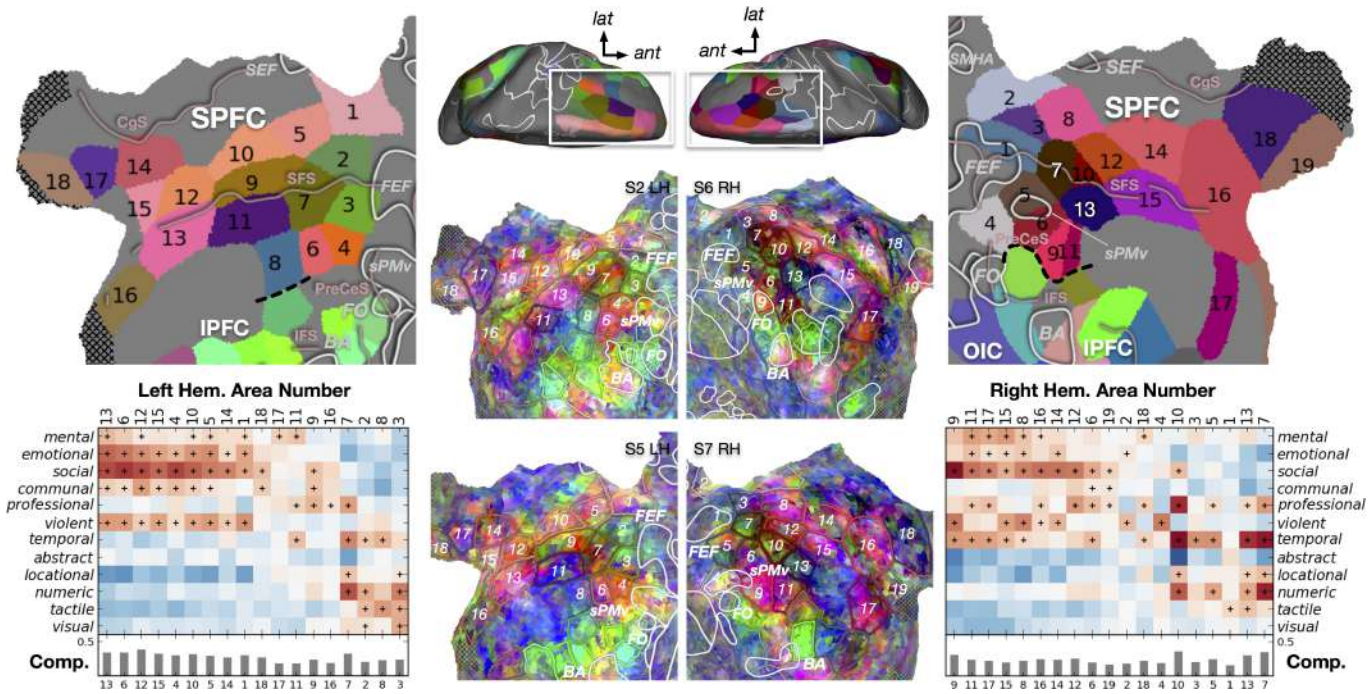
Extended Data Figure 5 | Comparison of PrAGMATiC models fit with different initial conditions. As with many clustering algorithms, PrAGMATiC optimizes a non-convex objective function and so can find many potential locally optimal solutions. To reduce the effect of non-convexity on our results, we re-fit the model ten times (each time with a different random initialization), and then selected the model fit that yielded the best likelihood (that is, performance on the training set) as the PrAGMATiC atlas (Fig. 3). Here we show the PrAGMATiC atlas (top) and the second best model out of the ten that were estimated (bottom). The parcellations given by these two models are very similar. However, there are a few differences, which illustrate uncertainty in the model. Some of these differences are due to statistical thresholding: a few areas that were found to be significantly semantically selective in the best model are missing in the alternative model (see left medial prefrontal cortex), and some significant areas in the alternate model are missing from the best model (left ventral occipital cortex). Other differences suggest alternative parcellations for a few regions, where, for example, the same region of cortex is parcellated into three areas in the best model and four areas in the alternative model. Yet it is clear that none of the differences between these two models are sufficient to change any of the interpretations given in the main text.



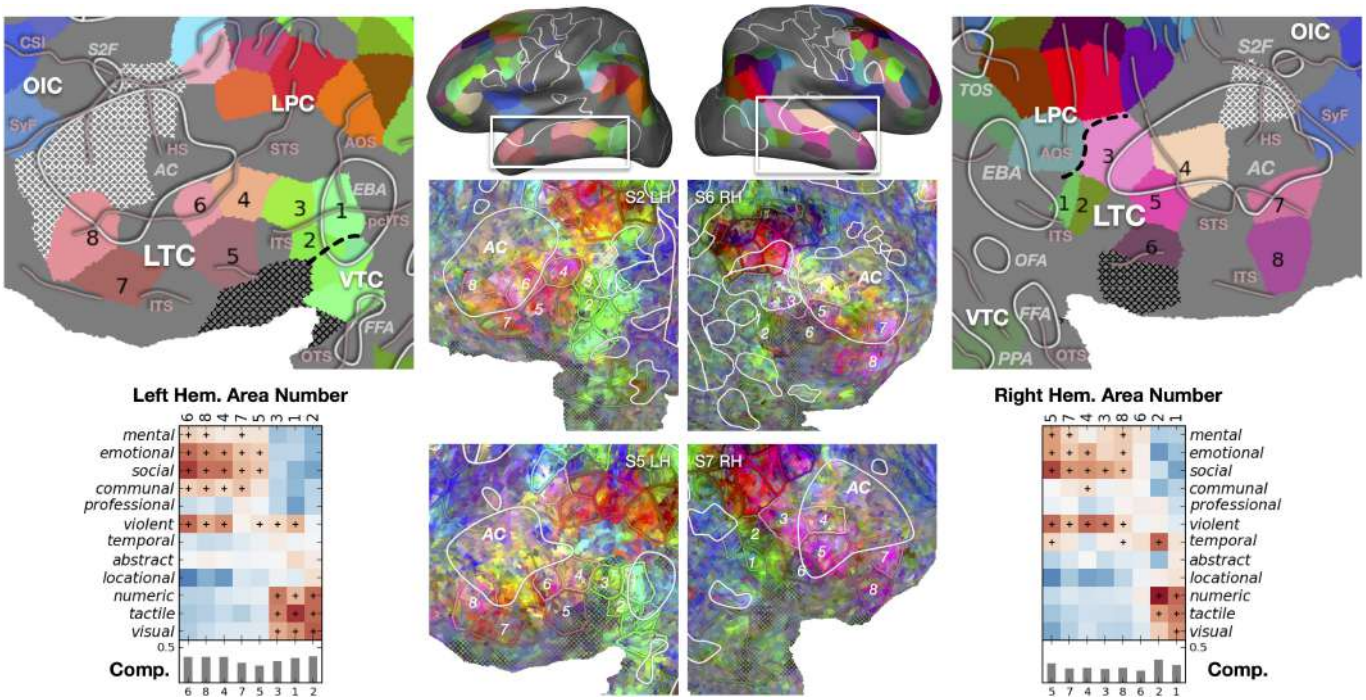
Extended Data Figure 6 | Semantic atlas for the LPC. The PrAGMATiC atlas divides the LPC into 15 areas in the left hemisphere and 13 areas in the right. Here we show the atlas for each hemisphere (top left and right), three-dimensional brains indicating the location of the LPC (top middle), individual maps for two subjects in each hemisphere (bottom middle), and the average predicted response of each area to the 12 semantic categories identified earlier (responses consistently greater than zero across subjects are marked with a plus) (bottom left and right). Bars show how completely this 12-category interpretation captures the average semantic model in each area. The LPC appears to be organized around the angular gyrus (AG), with a core that is selective for social, emotional and mental concepts (L6, 7, 9, 11; R5, 7) and a periphery that is selective for visual, tactile and numeric concepts (L2, 4, 5, 8, 10, 15; R6, 11).



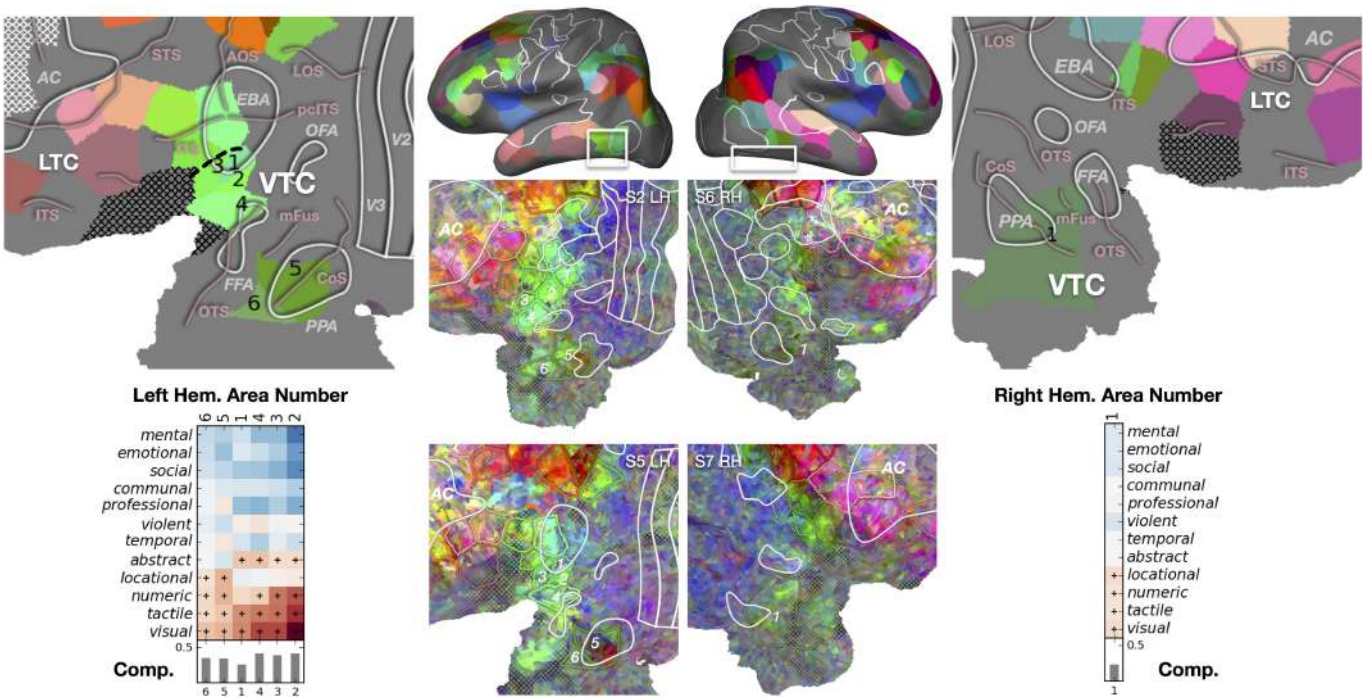
Extended Data Figure 7 | Semantic atlas for the MPC. The PrAGMATiC atlas divides the MPC into 14 areas in the left hemisphere and 10 areas in the right. Here we show the atlas for each hemisphere (top left and right), three-dimensional brains indicating the location of the MPC (top middle), individual maps for two subjects in each hemisphere (bottom middle), and the average predicted response of each area to the 12 semantic categories identified earlier (responses consistently greater than zero across subjects are marked with a plus) (bottom left and right). Bars show how completely the 12-category interpretation captures the average semantic model in each area. Like the LPC, the MPC appears to be organized around a core group of areas that are selective for social and mental concepts (L6, 8, 10; R6, 7). Dorsolateral MPC areas (L2, 4; R1) are selective for visual and tactile concepts. Anterior dorsal areas (L5, 9; R4, 9) are selective for temporal concepts. Ventral areas (L11, 12, 14; R8) are selective for professional, temporal and locational concepts. Just above the retrosplenial cortex one distinct area in each hemisphere is selective for mental, professional and temporal concepts (L7; R3). Overall, the right MPC responds more than the left MPC to mental concepts.



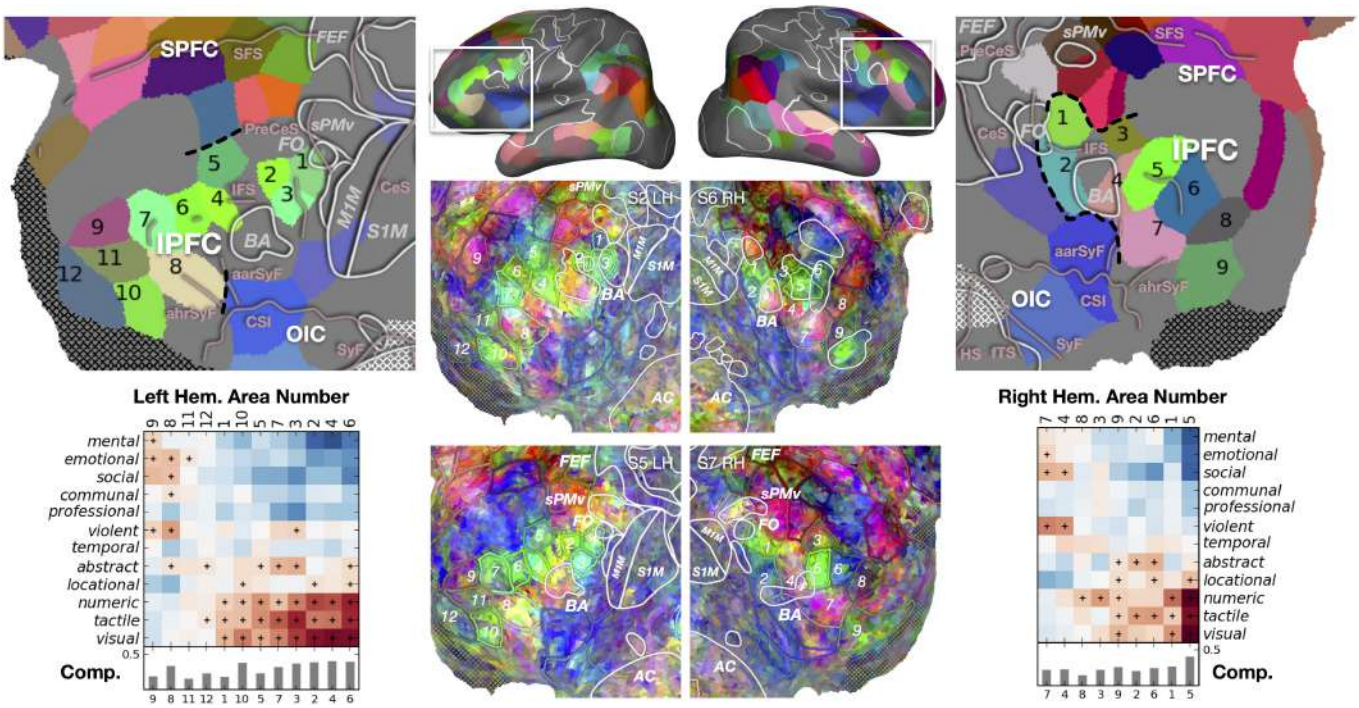
Extended Data Figure 8 | Semantic atlas for the SPFC. The PrAGMATiC atlas divides the SPFC into 18 areas in the left hemisphere and 19 areas in the right. Here we show the atlas for each hemisphere (top left and right), three-dimensional brains indicating the location of the SPFC (top middle), individual maps for two subjects in each hemisphere (bottom middle), and the average response of each area in the atlas to the 12 semantic categories identified earlier (responses consistently greater than zero across subjects are marked with a plus) (bottom left and right). Bars show how completely the 12-category interpretation captures the average semantic model in each area. The organization in the SPFC seems to follow the long rostro-caudal sulci and gyri of the dorsal frontal lobe. Posterior–lateral SPFC areas (L4, 6; R6, 9, 11) are selective for social, emotional, communal and violent concepts. Posterior superior frontal sulcus areas (L2, 3, 7, 8; R1, 5, 7) are selective for visual, tactile and numeric concepts. The superior frontal gyrus contains a long strip of areas (L1, 5, 10, 12–15; R8, 12, 14–16) selective for social, emotional, communal and violent concepts.



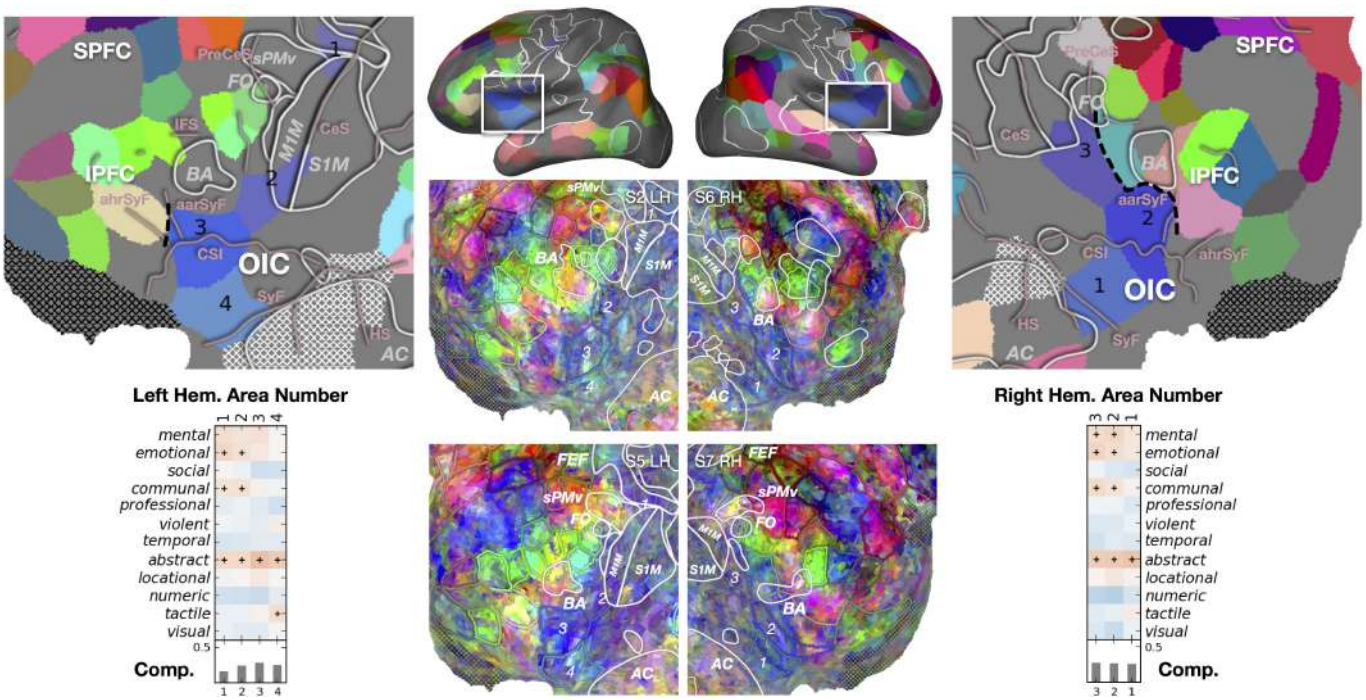
Extended Data Figure 9 | Semantic atlas for the LTC. The PrAGMATiC atlas divides the LTC into eight areas in both the left and right hemispheres. Here we show the atlas for each hemisphere (top left and right), three-dimensional brains indicating the location of the LTC (top middle), individual maps for two subjects in each hemisphere (bottom middle), and the average response of each area in the atlas to the 12 semantic categories identified earlier (responses consistently greater than zero across subjects are marked with a plus) (bottom left and right). Bars show how completely the 12-category interpretation captures the average semantic model in each area. Anterior LTC areas (L4–8; R3–8) are selective for social, emotional, mental and violent concepts. Posterior LTC areas (L1–3; R1–2) are selective for numeric, tactile and visual concepts.



Extended Data Figure 10 | Semantic atlas for the VTC. The PrAGMATiC atlas divides the VTC into six areas in the left hemisphere and one area in the right. Here we show the atlas for each hemisphere (top left and right), three-dimensional brains indicating the location of the VTC (top middle), individual maps for two subjects in each hemisphere (bottom middle), and the average response of each area in the atlas to the 12 semantic categories identified earlier (responses consistently greater than zero across subjects are marked with a plus) (bottom left and right). Bars show how completely the 12-category interpretation captures the average semantic model in each area. The VTC is relatively homogeneous: all areas are selective for numeric, tactile and visual concepts. Left VTC areas close to the parahippocampal place area (PPA) are also selective for locational concepts (L5–6).



Extended Data Figure 11 | Semantic atlas for the IPFC. The PrAGMATiC atlas divides the IPFC into 12 areas in the left hemisphere and 9 areas in the right. Here we show the atlas for each hemisphere (top left and right), three-dimensional brains indicating the location of the IPFC (top middle), individual maps for two subjects in each hemisphere (bottom middle), and the average response of each area in the atlas to the 12 semantic categories identified earlier (responses consistently greater than zero across subjects are marked with a plus) (bottom left and right). Bars show how completely the 12-category interpretation captures the average semantic model in each area. Posterior IPFC areas in the precentral sulcus (L1–3; R1, 2) are selective for visual, tactile and numeric concepts. Areas on the inferior frontal gyrus (L8; R4, 7) are selective for social and violent concepts. Areas in the inferior frontal sulcus and anterior middle frontal gyrus (L4–7; R5–6) are selective for visual, tactile and numeric concepts. Areas in the orbitofrontal sulci (L10; R9) are also selective for visual, tactile, numeric and locational concepts.



Extended Data Figure 12 | Semantic atlas for the opercular and insular cortex. The PrAGMATiC atlas divides the opercular and insular cortex (OIC) into four areas in the left hemisphere and three areas in the right. Here we show the atlas for each hemisphere (top left and right), three-dimensional brains indicating the location of the OIC (top middle), individual maps for two subjects in each hemisphere (bottom middle), and the average response of each area in the atlas to the 12 semantic categories identified earlier (responses consistently greater than zero across subjects are marked with a plus) (bottom left and right). Bars show how completely the 12-category interpretation captures the average semantic model in each area. These areas are homogeneously selective for abstract concepts, with more posterior and superior areas also responding to emotional, communal and mental concepts.