phonemes in English was spoken over time. This model was constructed from the phoneme representation of the stories: the lists of phoneme-time pairs *(P, t)* were re-arranged into 39 lists, each of which contains only the times of a single phoneme. These lists of times were then downsampled to the fMRI acquisition rate.

**9. Noise-ceiling correction.** While the correlation between predicted response and actual mean response is an appropriate metric for assessing significance, it is biased downward due to noise in the validation data (David & Gallant, 2005; Hsu, Borst, & Theunissen, 2004; Sahani & Linden, 2003). This is because the actual mean response is calculated using a finite number of repetitions (in this case 2) and thus it contains residual noise in addition to signal. This noise level is likely to vary across voxels due to vascularization and magnetic field inhomogeneity. For the corrected correlation flatmaps shown in Extended Data Figure 1, we accounted for noise in the validation data using the method developed in (Hsu et al., 2004). In this method the raw correlation is divided by the expected maximum possible model correlation (called the *noise ceiling*) for each voxel. For very noisy voxels, however, this method led to divergent correlation estimates. To correct this issue we limited voxel noise ceilings to be above some value $k$. For $k$=1, the estimated actual correlation is the observed correlation between response and prediction, and for $k$=0 the estimated actual correlation is the original divergent estimate. We used $k$=0.0966, which is the $p$<0.05 significance threshold for the correlation of two gaussian variables with the same length as our validation story.

**10. Significance testing of semantic principal components.** If there is no structured semantic space underlying the true model weights (i.e. the weights for each voxel are independent from the other voxels) then the PCs of the estimated model weights will be identical to the PCs of the stimulus matrix, which contains the semantic feature representations of each 2-second segment of the stories. This bias in the estimated weight PCs is due to the regularized regression procedure used here, which trades a small increase in bias for a large decrease in error (Hoerl & Kennard, 2012). Thus in order to appropriately evaluate statistical significance of the estimated model weight PCs we compared them to the PCs of the stories. This significance criterion helps ensure that the semantic structure that we observe in the PCs is due primarily to the fMRI data and not the statistics of the stories. We first tested whether each individual-subject model weight PC accounted for more variance than would be expected by chance. To find confidence intervals on the variance accounted for by each PC we bootstrapped the model weight PCA by sampling with replacement from the voxel population 1000 times. Similarly, confidence intervals on the variance in model weights accounted for by each story PC were obtained by bootstrapping the story PCA 1000 times. One potential issue with directly comparing the variance accounted for by an individual-subject PC and the correspondingly numbered story PC (i.e. comparing the first subject PC with the first story PC) is that the same PCs might appear in both analyses but in a different order. To account for this issue we re-ordered the first 20 story PCs to maximize their correspondence to the first 20 subject PCs using the Gale-Shapley stable marriage algorithm.

The amount of variance accounted for in the model weights by each of the model weight PCs (orange lines) and story PCs (gray lines) is shown in Extended Data Figure 2, along with error bars denoting 99% confidence intervals. To test the hypothesis that a model weight PC accounts for more variance than the corresponding story PC we counted the number of times in the 1000 bootstrap samples that the story PC accounted for more variance than the model weight PC. The null hypothesis for this analysis is that the story PC and the model weight PC account for the same amount of variance. We rejected the null hypothesis if the story PC never accounted for more variance than the voxel weight PC across the 1000 bootstrap samples (corresponding to $p$<0.001).

Because lower-variance PCs are more sensitive to noise and thus more likely to yield false positives, we tested the PCs sequentially and stopped testing after encountering the first non-significant PC. This procedure revealed that subject S1 has 6 significant individual-subject PCs, S2 has 8 significant PCs, S3 has 4 significant PCs, S4 has 6 significant PCs, S5 has 7 significant PCs, S6 has 6 significant PCs, and S7 has 4 significant PCs.

Next we tested PCs constructed using combined data from many subjects. For each subject we constructed a set of group PCs using combined data from the other six subjects, leaving out the selected subject. For example, to test subject S1 we performed PCA on combined model weights from subjects S2-S7. We then computed the amount of variance accounted for in the model weights for the left out subject by each of the group PCs. As with the individual subject PCs and story PCs, confidence intervals on the variance explained by the group PCs were found using the bootstrap. The amount of variance accounted for in the model weights by each of the group PCs (blue lines) is shown in Extended Data Figure 2, along with error bars denoting 99% confidence intervals.

We then tested whether each group PC explained more variance than the corresponding story PC (again re-ordered using the Gale-Shapley stable marriage algorithm) using the statistical procedure described above. We found that subject S5 was significantly explained by 6 group PCs, subjects S1 and S3 were significantly explained by 5 group PCs, subjects S4, S6, and S7 were significantly explained by 4 group PCs, and subject S3 was significantly explained by 3 group PCs (Extended Data Figure 2).

**11. Semantic word cluster analysis.** Cluster analysis was used to create interpretable features in the semantic space. First all 10,470 words in the semantic feature space were projected into the 4-dimensional common semantic PC space. Then an iterative, robust convex hull estimation procedure was used to find the most important words in this space. At each iteration, 80% of the 10,470 words were selected at random, and then their convex hull was found in the 4-D semantic space. This was repeated 100 times. The set of all words that appeared on the convex hull in at least one iteration was then found. These 458 words were then clustered in the 4-D space using the *k*-means implementation in scikit-learn (Pedregosa et al., 2011). To select the number of clusters we computed the fraction of variance that the clusters collectively explained in the mean semantic model for each of the significant semantic areas identified by the PrAGMATiC atlas. Then we selected the smallest number of clusters that would account for at least 10% of the variance in each PrAGMATiC area, which was 12. To maximize cluster stability we repeated the *k*-means clustering 100 times and selected the model with the highest average variance explained across the PrAGMATiC areas. Within each *k*-means repetition the clustering model was initialized 100 times using *k*-means++. Labels were assigned to the clusters manually by inspecting the words that appeared in each cluster (Supplementary Table 2). For alternate label assignment methods see Supplementary Data 5.

**12. PrAGMATiC details.** The PrAGMATiC algorithm assumes that the cortex of each subject is tiled with convex functional areas, and that all locations within each area have the same tuning within the 4-dimensional semantic space. The location of each area is determined by the location of its centroid, which is a single point on the cortical surface. The location of each centroid depends on the locations of a few neighboring centroids and the locations of some known landmarks, which are identified separately in each subject. The functional selectivity of each area is determined by its mean functional value. The mean functional value for area $i$ is called $M_i$.

This model is instantiated for each subject as a two-layer Bayesian network, with one visible layer and one hidden layer. The visible layer units are vertices on the cortical surface mesh. Each vertex is associated with a $D$-vector of observed functional values. The vector of observed values for visible unit $l$ in subject $s$ is called $v_{ls}^{obs} \in \mathcal{R}^D$, and the collection of all visible units in a subject is called $V_s^{obs}$. We assume that all visible units are independent of each other, given the hidden layer units.

The hidden layer units are the locations of the area centroids. The location of centroid $i$ in subject $s$ is called $h_{is}$, and the collection of all hidden units in subject $s$ is called $H_s$.

As a generative model, PrAGMATiC must be able to generate samples from the distribution of visible unit vectors. To sample $v_{ls}$ we first find the index of the nearest area centroid on the cortical surface, $c(H, l, s)$. Then we look up the mean associated with that centroid, $M_{c(H,l,s)}$. Finally we draw a sample from a multivariate Gaussian distribution with spherical variance: $v_{ls} \sim \mathcal{N}\left(M_{c(H,l,s)}, \sigma_V^2 I_D\right)$

The probability distribution over locations of the hidden units is modeled using a physical analogy to a system of springs. The ideal length of the spring connecting units $i$ and $j$ is called $L_{ij}$.

The full probability distribution for PrAGMATiC is written:

$$\begin{aligned} P(V, H; M, L) &= P(H; L) P(V|H; M) \\ &= \left[Z_H(L)^{-1} e^{-E(H;L)}\right] \left[Z_V^{-1} e^{-E(V|H;M)}\right] \end{aligned}$$

The distribution over arrangements of the hidden layer units, $H$, is modeled using a Boltzmann distribution with the following energy function:

$$E(H; L) = \frac{\beta}{2} \sum_{i,j,s} (d_{ijs} - L_{ij})^2$$

Here $d_{ijs}$ is the geodesic distance across the cortical surface between hidden layer units $i$ and $j$ in subject $s$. This distance is computed using a heat-based approximation to the exact geodesic distance (Crane et al. 2012). This energy function is exactly the sum over the spring potential energy for all spring connections in the model. The constant $\beta$ determines the temperature of the spring system. The normalizing constant for $P(H; L)$ depends on the value of $L$, and is written here as $Z_H(L)$.

The distribution over visible unit values is multivariate Gaussian with equal variance in all dimensions and zero covariance, but for consistency we write it as an energy-based model. The energy function for the visible units is:

$$E(V|H; M) = \frac{\sigma_V^{-2}}{2} \sum_{l,s} \left(v_{ls} - M_{H(s,l)}\right)^2$$

Here $M_{H(s,l)}$ is the mean functional value for the closest hidden layer unit (by geodesic distance across the cortical surface) in the arrangement $H$ to visible layer unit $l$ in subject $s$. The constant $\sigma_V$ is the standard deviation of the Gaussian. The normalizing constant for $P(V|H; M)$ depends on $\sigma_V$, but not on any of the learned parameters (because this is a Gaussian distribution its normalizing constant is known).

We use maximum likelihood estimation (MLE) to learn $L$ and $M$ based on observed visible unit data, $V^{obs}$

$V^{obs}$. For the spring lengths, the average log likelihood given the observed data is written:

$$\mathscr{L}(L; V^{obs}) = \frac{1}{N}\sum_s \log P(V_s^{obs}; L)$$

Here $N$ is the number of subjects and $s$ is an index across subjects.

Then we differentiate with respect to $L$ to find:

$$\begin{aligned}\frac{\partial \mathscr{L}(L; V^{obs})}{\partial L} &= \left(\frac{\partial Z_H(L)}{\partial L}\right)^{-1} Z_H(L) \\ &\quad - \frac{1}{N}\sum_s \frac{\sum_h P(V_s^{obs}|h; M)P(h; L)\frac{\partial E(h;L)}{\partial L}}{P(V_s^{obs}; L, M)}\end{aligned}$$

Where the total probability of the observed data given the parameters, $P(V_s^{obs}; L, M)$ is equal to the expectation over $H$:

$$P(V_s^{obs}; L, M) = \sum_h P(V_s^{obs}|h; M)P(h; L)$$

The first part of the gradient, which involves the normalization constant $Z_H$, can be written as:

$$\left(\frac{\partial Z_H(L)}{\partial L}\right)^{-1} Z_H(L) = \frac{1}{N}\sum_{s,h} P(h; L)\frac{\partial E(h; L)}{\partial L}$$

or simply as the expectation of the gradient over $H$:

$$\left(\frac{\partial Z_H(L)}{\partial L}\right)^{-1} Z_H(L) = \left\langle \frac{\partial E(H; L)}{\partial L} \right\rangle_H$$

Note that the entire gradient could be written more simply as the Boltzmann learning rule from Ackley, Hinton, & Sejnowski, 1985:

$$\frac{\partial \mathscr{L}(L; V^{obs})}{\partial L} = \left\langle \frac{\partial E(H; L)}{\partial L} \right\rangle_H - \left\langle \frac{\partial E(H; L)}{\partial L} \right\rangle_{H|V^{obs}}$$

However, to make an essential approximation we retain the earlier formulation. This gradient is impossible to compute exactly because it requires integrating over all possible $H$. Therefore we approximate the gradient using only a small number of samples from $P(H; L)$. These samples are obtained using Gibbs sampling, wherein the location of each hidden unit is update sequentially according to the conditional distribution $P(h_{is}|H_{/is}; L)$. This procedure is used to obtain $J$ samples of $H$, which are denoted $\tilde{H}^j$ with $j = 1...J$. We then use these samples to approximate the integral and expectation over $H$. The gradient function is then rewritten as:

$$\frac{\partial \mathscr{L}(L;V^{obs})}{\partial L} = \frac{1}{NJ}\sum_{j,s}\frac{\partial E(\tilde{H}_s^j;L)}{\partial L} - \frac{1}{N}\sum_{s}\frac{\sum_j P(V_s^{obs}|\tilde{H}_s^j;M)\frac{\partial E(\tilde{H}_s^j;L)}{\partial L}}{\sum_j P(V_s^{obs}|\tilde{H}_s^j;M)}$$

This function shows that the likelihood gradient is equal to the difference between the average energy gradient across all $J$ samples (the first term) and a weighted average energy gradient (the second term), where the weights are proportional to the probability of the observed data $V^{obs}$ given the sampled $H$.

To compute the energy gradient for each sample we differentiate the energy function with respect to each element of $L$, giving:

$$\frac{\partial E(H;L)}{\partial L_{ij}} = \frac{\beta}{2}\sum_{s}(L_{ij} - d_{ijs})$$

The gradient for $M$ is slightly different because the normalization constant $Z_V$ does not depend on $M$ (as the normalization constant of a Gaussian does not depend on the mean). Thus it has a simpler expression:

$$\frac{\partial \mathscr{L}(M;V^{obs})}{\partial M} = -\frac{1}{N}\sum_{s}\frac{\sum_j P(V_s^{obs}|\tilde{H}_s^j;M)\frac{\partial E(V^{obs}|\tilde{H}_s^j;M)}{\partial M}}{\sum_j P(V_s^{obs}|\tilde{H}_s^j;M)}$$

And the energy gradient for the mean of area $i$, $\partial E_V/\partial M_i$ is:

$$\frac{\partial E(V|H;M_i)}{\partial M_i} = \sigma_V^{-2}\sum_{l,s;H(s,l)=i}(V_{ls} - M_i)$$

Where the sum is taken only over the visible units $l, s$ for which the closest hidden unit in the arrangement, $c(H, l, s)$ is $i$.

To obtain high quality, independent samples of $H$ we maintain $J$ parallel Markov chains for each of the $N$ subjects. At each learning step we perform one Gibbs sweep through each of the Markov chains. That is, at step $t$ in chain $j$ and subject $s$ we draw the sample:

$$\tilde{H}_s^{j,t} \sim P(H_s|\tilde{H}_s^{j,t-1};L^t)$$

For each of the $J$ samples we compute the energy gradients for $L$ and $M$, as well as the likelihood of the observed data $P(V_s^{obs}|\tilde{H}_s^{j,t};M^t)$. Then we compute the average gradients and the weighted average gradient according to the data likelihoods. Finally we update $L$ and $M$ by taking a small step down these gradients:

$$L^{t+1} = L^t - \epsilon\frac{\partial \mathscr{L}(L^t;V^{obs})}{\partial L^t}$$

$$M^{t+1} = M^t - \epsilon\frac{\partial \mathscr{L}(M^t;V^{obs})}{\partial M^t}$$

The learning rate, $\epsilon$, is set on each step so that the largest change in any spring length is no more than

2mm and the largest change in any mean functional value is no more than 0.025 standard deviations.

The hyperparameters $\beta$ and $\sigma_V$ affect learning speed, but they do not directly affect the learned parameters (except by virtue of poor approximation). The inverse spring temperature, $\beta$, determines how stiff or floppy the springs are. If the inverse temperature is very high then the springs will be very stiff, samples of $H$ will be highly correlated across iterations, and the quality of the gradient steps will suffer. If it is very low then the springs will be very floppy, samples of $H$ will be highly random, and the quality of the gradient steps will also suffer.

If $\sigma_V$ is very low, then one sample from $H$ will always yield much higher likelihood of $V^{obs}$ than the others, and the weighted average of the gradients across samples will become the difference between the best sample and the other samples. If $\sigma_V$ is very high, then the likelihood of $V^{obs}$ will be very similar for all samples, and the weighted average of the gradients will be almost identical to the simple average across gradients.

Note further that these hyperparameters interact with each other. If $\beta$ is very high, then almost all samples from $H$ will be close to the $H$ that minimizes the total spring energy. Because the samples will be more similar, the likelihoods will also be more similar, and the weighted average of the gradients will again be similar to the simple average. The hyperparameters also interact with the number of areas in the model.

Rather than tuning these parameters directly, we select desired levels of entropy for $P_h = P(h_{is}|H_{/is}; L)$ and $P_V = P(V_s^{obs}|\tilde{H}_s^j; M)$. If the average entropy of $P_h$ is lower than the target value then we lower $\beta$ to make the springs more floppy; if it is higher than the target value then we raise $\beta$ to make the springs stiffer. Similarly, if the entropy of $P_V$ is lower than the target value then we raise $\sigma_V$; if it is higher than the target value then we lower $\sigma_V$.

High entropy keeps the model from falling into local minima, but also keeps the model from finding very high likelihood solutions. Conversely, low entropy allows the model to find high likelihood solutions, but also makes it more likely to fall into local minima. To take advantage of both low and high entropy states we use an annealing approach, where the entropy target for $P_h$ is high at the beginning of learning, but then is gradually lowered throughout the learning process. This makes the Markov chain take larger, more uncorrelated steps at the beginning of learning, but smaller steps at the end.

In practice the algorithm as written above tends to converge when the numbers of areas and subjects are both small. If these numbers become large, however, (e.g. 128+ clusters and 5+ subjects) the algorithm becomes less stable. When this occurs, the model tends to prioritize minimum energy solutions over maximum probability solutions. That is, the model tries to minimize the total spring energy across all the subjects at the cost of poorly explaining the data. This often causes all the areas to bunch up as far as possible from any known landmarks. These effects are exacerbated when $\beta$ is high.

We believe that this problem is caused by bias in drawing samples of $H$. When the spring temperature is low, all samples from $P(H; L)$ will be very close to the minimum spring energy state (i.e. the arrangement that minimizes $E(H; L)$). If the minimum spring energy state is far from the maximum probability state (i.e. the arrangement that maximizes $P(V^{obs}|H)$), then this could bias the gradient steps such that the likelihood decreases over time.

One solution to this problem is to increase $J$, the number of samples that are drawn for each subject at each step of the learning algorithm. However, this is expensive, as run time is linear in $J$. An alternative solution that incurs almost no additional cost is to add a small perturbation to the ideal spring lengths that is specifically tailored to each subject. That is, we replace the global spring length parameters $L$ with $L + \Lambda_s$. Now the domain of possible minimum spring energy states is much larger, and thus it is much more likely that the maximum probability state also has low or minimum spring energy. To ensure that $\Lambda_s$ stays small and does not capture differences that are common across subjects we set the learning rate for $\Lambda_s$ to 1/10 of the learning rate for $L$.

With this new term, the full probability distribution becomes:

$$\begin{aligned} P(V, H; M, L, \Lambda) &= P(H; L, \Lambda) P(V|H; M) \\ &= \left[ Z_H(L, \Lambda)^{-1} e^{-E(H;L,\Lambda)} \right] \left[ Z_V^{-1} e^{-E(V|H;M)} \right] \end{aligned}$$

And the spring energy function becomes:

$$E(H; L, \Lambda) = \frac{\beta}{2} \sum_{i,j,s} (d_{ijs} - L_{ij} - \Lambda_{ijs})^2$$

The energy gradient for $\Lambda$ is very similar to that for $L$:

$$\frac{\partial E(H; L, \Lambda)}{\partial \Lambda_{ijs}} = \frac{\beta}{2} (L_{ij} + \Lambda_{ijs} - d_{ijs})$$

PrAGMATiC optimizes a non-convex objective function and so can find many potential locally optimal solutions. (This is a common problem in clustering algorithms.) This also means that the resulting model is sensitive to initial conditions and to the random seed that is used while sampling new model states during learning. Here we use two methods to maximize the chance of finding the optimal global solution. First, we use an annealing process, lowering the temperature parameter slowly over the course of learning. Second, we re-estimate the model for each hemisphere 10 times, using a different random initialization and random seed each time. From these various estimates we take the model that has the highest likelihood to be the canonical model. Extended Data Figure 5 shows the model with the highest likelihood, and the model with the second-highest likelihood. The most likely and second-most likely functional parcellations are quite similar, though there are some small local differences that reflect statistical thresholding and the influence of initial conditions. Thus, our implementation of the PrAGMATiC algorithm seems to produce reliable and stable estimates of functional parcellation.

**13. PrAGMATiC atlas.** To determine how many total areas should be used in the PrAGMATiC atlas we used a cross-validation procedure. We estimated PrAGMATiC models with different numbers of areas (ranging from 8 to 384) and data from six of the seven subjects. We computed the average semantic voxel-wise model weight vector (including all four delays) in each area for each of the six subjects included in the PrAGMATiC model. This was done by projecting the weight vectors into vertex space using the *line-nearest* scheme in pycortex and then averaging across all the vertices within each area. Then those vectors were averaged together across subjects to obtain a single estimated weight vector for each area.

Next we used the estimated PrAGMATiC model to parcellate cortex in the seventh subject, generating a predicted parcellation based only on the locations of the functional landmarks. This was done by loading the spring lengths from the estimated PrAGMATiC model and then resampling the area centroid locations in the seventh subject for 100 iterations, during which the inverse temperature (beta) was gradually increased from 0.05 to 37. At this point each vertex in the seventh subject is assigned to a single area in the PrAGMATiC parcellation.

Next we tested how well the average weight vector for each PrAGMATiC area from the other six subjects could predict responses in the seventh subject. We projected BOLD responses to the validation story (which was not used for voxel-wise model estimation) for the seventh subject into vertex space using the *line-nearest* scheme in pycortex. Then for each vertex we used the average weight vector assigned to its area based on the other six subjects to predict BOLD responses to the validation story in the seventh subject. Finally we computed the correlation between the predicted and actual BOLD responses to obtain a measure of model prediction performance. These correlation values were averaged across all vertices in each hemisphere in each subject to obtain the PrAGMATiC prediction values shown in Figure 3B. This procedure was repeated three times while holding out each of the seven subjects separately.

To select the best number of areas for each hemisphere based on these correlation values, we performed a linear mixed-effects ANOVA using the lmer package in R with each number of areas as a factor level in a fixed effect and subjects as random effects. This showed that correlation was significantly different for different numbers of areas. Next we performed pairwise post hoc tests comparing mean correlations across numbers of areas using the multcomp package in R. Resulting *p*-values were corrected for multiple comparisons using FDR. Finally, we selected the smallest number of areas for which the mean correlation was not significantly different (q(FDR)>0.01) from the mean correlation for any larger number of areas. For the left hemisphere, this required 192 areas. For the right, this required 128 areas.

The final PrAGMATiC atlas was based on data from all seven subjects, unlike the cross-validated models described above. To identify semantically selective areas in the atlas we tested whether the average semantic model in each area performed significantly better than the average low-level model. This was done using a similar procedure to that outlined above. The average semantic model was computed for each area (including data from all seven subjects), as was the average low-level model for each area. Predictions for each vertex in each area were computed, as above, for both semantic and low-level models. Then a bootstrap procedure was used to estimate a distribution of correlation values for each vertex under each model by resampling the 290 time points that were used to compute the correlation, with replacement, 1000 times. These bootstrap correlations were averaged across each area and across subjects. Finally the average correlations for the semantic and low-level models were compared for each bootstrap sample. The *p*-value for the significance test was computed as the fraction of samples where the average semantic model correlation for an area was less than the average low-level model correlation. This *p*-value is 1.0 for areas where the low-level model is always better, 0.5 for areas where both models are about the same, and 0.0 for areas where the semantic model is always better. These *p*-values were corrected for multiple comparisons using FDR. Corrected *p*-value thresholds were chosen based on the total number of areas. For example, in the left hemisphere the threshold $p<1/192$ should limit the number of non-semantic areas that are considered semantic to fewer than one. Only areas where the semantic model performed significantly better than the low-level model according to this test are shown in Figure 3 and Extended Data Figures 6-12.

To describe the semantic selectivity of areas in the PrAGMATiC atlas we predicted the average response of each area to each of the 12 semantic categories identified earlier (Figure 2A). This was done by computing the average semantic model weights for each area (as described above, but here averaging across delays) and then projecting those weights onto the average semantic vector for each of the 12 categories. These values are shown in Extended Data Figures 6-12. To determine whether each area was selective for a category we used a *t*-test to determine whether the estimated response to the category was consistently greater than zero across subjects (q(FDR)<0.05, FDR correction applied across areas within each region and across the 12 categories).

To order the 12 categories for display purposes (Extended Data Figures 6-12) we projected the category vectors onto semantic model weights for all areas in the left hemisphere, and then computed the correlation between these projections for each pair of categories. Then we used a traveling salesman solver to find a path through the 12 categories that maximized correlations between adjacent categories.

One issue with using the 12 semantic categories to describe semantic selectivity is that many semantic concepts or categories that might be represented in the brain will fall outside of those categories. This would lead the 12 category interpretation to be incomplete. To assess how completely the 12 category interpretation describes each area in the PrAGMATiC atlas we fit linear models that attempted to recreate the average semantic model weights for each area from a weighted combination of the semantic vectors for the 12 categories. Then we computed the fraction of variance in the average semantic model weights that was explained by this linear model. In the best-explained areas the 12 categories account for 40-50% of the variance in the average semantic model weights, while for poorly explained areas they can account for less than 15% of the variance. Areas with very low variance explained are incompletely described by the 12 categories, while areas with higher variance explained are well described by the 12 categories. The variance explained for each area is shown in bar plots in Extended Data Figures 6-12.

To determine whether left hemisphere or right hemisphere areas within any given region (such as the medial parietal cortex) were significantly more selective for any of the 12 semantic categories we used a *t*-test to compare predicted responses in all left hemisphere areas to all right hemisphere areas.

## Supplementary References

Ackley, D., Hinton, G., & Sejnowski, T. (1985). A learning algorithm for boltzmann machines. *Cognitive Science*, *9*(1), 147–169. doi:10.1016/S0364-0213(85)80012-4

Baldo, J. V, & Shimamura, a P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology*, *12*(2), 259–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9556772

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527–36. doi:10.1016/j.tics.2011.10.001

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex (New York, N.Y. : 1991)*, *19*(12), 2767–96. doi:10.1093/cercor/bhp055

Bookheimer, S. (2002). Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*, 151–88. doi:10.1146/annurev.neuro.25.112701.142946

Bowers, D., Verfaellie, M., Valenstein, E., & Heilman, K. M. (1988). Impaired Acquisition of Temporal Information Amnesia in Retrosplenial, *66*, 47–66.

Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, *1124*, 1–38. doi:10.1196/annals.1440.011

Courtney, S. M. (1998). An Area Specialized for Spatial Working Memory in Human Frontal Cortex. *Science*, *279*(5355), 1347–1351. doi:10.1126/science.279.5355.1347

Dale, A., Fischl, B., & Sereno, M. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, *194*, 179–194. Retrieved from ftp://132.183.202.158/pub/articles/1999/1999_-_Dale_et_al._-_NeuroImage.pdf

David, S. V., & Gallant, J. L. (2005). Predicting neuronal responses during natural vision. *Network: Computation in Neural Systems*, *16*(2-3), 239–260. doi:10.1080/09548980500464030

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science (New York, N.Y.)*, *293*(5539), 2470–3. doi:10.1126/science.1063414

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601. doi:10.1038/33402

Gao, J. S., Huth, A. G., Lescroart, M. D., & Gallant, J. L. (2015). Pycortex: an interactive surface visualizer for fMRI. *Frontiers in Neuroinformatics*, *9*(September), 1–12.

doi:10.3389/fninf.2015.00023

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, *8*(5), 393–402. doi:10.1038/nrn2113

Hoerl, A. E., & Kennard, R. W. (2012). Ridge Regression : Biased Estimation for Nonorthogonal Problems, *12*(1), 55–67.

Hsu, A., Borst, A., & Theunissen, F. (2004). Quantifying variability in neural responses and its application for the validation of model predictions. *Network: Computation in Neural Systems*, *15*(2), 91–109. doi:10.1088/0954-898X/15/2/002

Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., & Fiske, A. P. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *NeuroImage*, *21*(3), 1167–73. doi:10.1016/j.neuroimage.2003.11.013

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11516708

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. doi:10.1007/BF02289233

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). 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*, *17*(11), 4302–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9151747

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, a, … Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain : A Journal of Neurology*, *123 Pt 12*, 2512–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11099452

Levy, R., & Goldman-Rakic, P. S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. *Experimental Brain Research*, *133*(1), 23–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10933207

Linden, D. E. J., Oosterhof, N. N., Klein, C., & Downing, P. E. (2012). Mapping brain activation and information during category-specific visual working memory. *Journal of Neurophysiology*, *107*(2), 628–39. doi:10.1152/jn.00105.2011

Lüders, H., Lesser, R. P., Hahn, J., Dinner, D. S., Morris, H. H., Wyllie, E., & Godoy, J. (1991). Basal temporal language area, *1986*(February 1986), 743–754.

Lundstrom, B. (2003). Isolating the retrieval of imagined pictures during episodic memory: activation of the left precuneus and left prefrontal cortex. *NeuroImage*, *20*(4), 1934–1943.

doi:10.1016/j.neuroimage.2003.07.017

McDonald, C. R., Crosson, B., Valenstein, E., & Bowers, D. (2001). Verbal Encoding Deficits in a Patient with a Left Retrosplenial Lesion. *Neurocase*, *7*(5), 407–417. doi:10.1076/neur.7.5.407.16250

Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science (New York, N.Y.)*, *343*(6174), 1006–10. doi:10.1126/science.1245994

Mikolov, T., Yih, W., & Zweig, G. (2012). Linguistic Regularities in Continuous Space Word Representations.

Nakamura, K., Kawashima, R., Sato, N., Nakamura, a, Sugiura, M., Kato, T., … Zilles, K. (2000). Functional delineation of the human occipito-temporal areas related to face and scene processing. A PET study. *Brain : A Journal of Neurology*, *123 ( Pt 9*, 1903–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10960054

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. doi:10.1016/0028-3932(71)90067-4

Ono, M., Kubik, S., & Abernathey, C. D. (1990). *Atlas of the Cerebral Sulci*. G. Thieme Verlag. Retrieved from https://books.google.com/books?id=xroe986wtkEC

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peelen, M. V, Romagno, D., & Caramazza, A. (2012). Independent representations of verbs and actions in left lateral temporal cortex. *Journal of Cognitive Neuroscience*, *24*(10), 2096–107. doi:10.1162/jocn_a_00257

Price, C. J. (2010). The anatomy of language: a review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*, 62–88. doi:10.1111/j.1749-6632.2010.05444.x

Raichle, M. E., MacLeod, a M., Snyder, a Z., Powers, W. J., Gusnard, D. a, & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–82. doi:10.1073/pnas.98.2.676

Robinson, G., Blair, J., & Cipolotti, L. (1998). Dynamic aphasia : an inability to select between competing verbal responses ?, 77–89.

Sahani, M., & Linden, J. F. (2003). How Linear are Auditory Cortical Responses ? ∗.

Thompson-Schill, S. L., Swick, D., Farah, M. J., D'Esposito, M., Kan, I. P., & Knight, R. T. (1998). Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(26), 15855–60. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=28134&tool=pmcentrez&rendertype=abstract

Verstynen, T. D., & Deshpande, V. (2011). Using pulse oximetry to account for high and low frequency physiological artifacts in the BOLD signal. *NeuroImage*, *55*(4), 1633–44. doi:10.1016/j.neuroimage.2010.11.090

Visser, M., Jefferies, E., & Lambon Ralph, M. a. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, *22*(6), 1083–94. doi:10.1162/jocn.2009.21309

Wagner, A. D., Shannon, B. J., Kahn, I., & Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, *9*(9), 445–53. doi:10.1016/j.tics.2005.07.001

Warrington, E. K. (1975). The selective impairment of semantic memory. *Q. J. Exp. Psychol.*, *27*, 635–657. doi:10.1080/14640747508400525

doi:10.1038/nature17637



# Supplementary Tables

**1. Semantic categories that are over- or under-sampled in the story stimuli.** To determine which semantic categories are over- or under-sampled in the stories used in this experiment we compared the stories to a large text corpus (see Supplementary Data 1 for details). Semantic categories were found using Ward agglomerative clustering. For each category that was significantly over- or under-sampled in the stories, this table shows the frequency of that category in the corpus, the frequency of that category in the story stimuli, the ratio of those frequencies, the *p*-value of the difference in frequency, and a few sample words from that category.

## Under-sampled categories

| Corpus frequency | Stimulus frequency | Ratio ↓ | *p*-value | Sample words |
|---|---|---|---|---|
| 47.33/100k | 0.00/100k | inf:1 | 0.000029 | sinking stern boat diving sank |
| 38.08/100k | 0.00/100k | inf:1 | 0.000284 | mademoiselle dorrit mrs winkle squeers |
| 301.35/100k | 8.57/100k | 35.2:1 | 0.000000 | unit laboratory foundation project extend |
| 83.50/100k | 4.29/100k | 19.5:1 | 0.000000 | sailed boarded arriving aboard patrol |
| 184.44/100k | 12.86/100k | 14.3:1 | 0.000000 | admiral military artillery survivors disaster |
| 167.77/100k | 17.14/100k | 9.8:1 | 0.000000 | succession british crowned kingdom rebellion |
| 569.27/100k | 60.00/100k | 9.5:1 | 0.000000 | adapted originally major currently expanded |
| 85.37/100k | 12.86/100k | 6.6:1 | 0.000008 | entering temporarily reserve guidance aid |
| 240.43/100k | 38.57/100k | 6.2:1 | 0.000000 | organized nations affairs nation dispute |
| 102.32/100k | 17.14/100k | 6.0:1 | 0.000002 | pulse channels tracking via radios |
| 310.61/100k | 55.72/100k | 5.6:1 | 0.000000 | machine effectively devices thereby equipped |
| 465.40/100k | 90.00/100k | 5.2:1 | 0.000000 | woods meadows grove canyon eastern |
| 215.87/100k | 42.86/100k | 5.0:1 | 0.000000 | union commerce cities occupation towns |
| 244.63/100k | 51.43/100k | 4.8:1 | 0.000000 | instrument introduction purposes instruments traced |
| 135.24/100k | 30.00/100k | 4.5:1 | 0.000000 | threatened enemies destroy brutal threatening |
| 647.01/100k | 154.29/100k | 4.2:1 | 0.000000 | circumstances responsible altogether privilege citizens |
| 141.41/100k | 34.29/100k | 4.1:1 | 0.000000 | elegant elaborate females likeness mature |
| 508.09/100k | 132.86/100k | 3.8:1 | 0.000000 | headquarters halls stables city boulevard |
| 80.55/100k | 21.43/100k | 3.8:1 | 0.000311 | lord highness privy constable priest |
| 218.99/100k | 60.00/100k | 3.6:1 | 0.000000 | peril disguise spared summoned siege |
| 136.92/100k | 38.57/100k | 3.5:1 | 0.000003 | collapse reactor atomic shifting observe |
| 181.23/100k | 51.43/100k | 3.5:1 | 0.000000 | miller owens barton maxwell marshall |
| 686.19/100k | 197.15/100k | 3.5:1 | 0.000000 | o w d span h |
| 118.42/100k | 34.29/100k | 3.5:1 | 0.000023 | cops prison unaware suspicion pleaded |
| 250.62/100k | 72.86/100k | 3.4:1 | 0.000000 | resigned rival representative presidential alliance |
| 73.45/100k | 21.43/100k | 3.4:1 | 0.001454 | possum elephant wolves turtle butterfly |
| 219.92/100k | 64.29/100k | 3.4:1 | 0.000000 | estimate similarly object y scale |
| 251.57/100k | 77.14/100k | 3.3:1 | 0.000000 | increase increasing absorbed owing million |
| 264.71/100k | 81.43/100k | 3.3:1 | 0.000000 | fail logic argue opinions suspect |
| 188.16/100k | 60.00/100k | 3.1:1 | 0.000000 | preserve suitable ideal unique lacked |
| 303.19/100k | 98.57/100k | 3.1:1 | 0.000000 | represents borne contained representing refer |
| 774.81/100k | 257.15/100k | 3.0:1 | 0.000000 | tree pavement feet rolling ground |
| 153.38/100k | 51.43/100k | 3.0:1 | 0.000008 | lethal survival beings power fatal |
| 126.54/100k | 42.86/100k | 3.0:1 | 0.000064 | trade proprietor owned owner co |
| 113.65/100k | 38.57/100k | 2.9:1 | 0.000190 | metallic patch uneven spot canvas |

| Corpus frequency | Stimulus frequency | Ratio ↓ | *p*-value | Sample words |
|---|---|---|---|---|
| 86.89/100k | 30.00/100k | 2.9:1 | 0.001181 | drying melt container freezing temps |
| 95.51/100k | 34.29/100k | 2.8:1 | 0.000928 | spiders herd snakes spider insects |
| 105.91/100k | 38.57/100k | 2.7:1 | 0.000556 | spoke uttered spoken mute word |
| 176.43/100k | 64.29/100k | 2.7:1 | 0.000005 | wells detached divided remainder prior |
| 175.83/100k | 64.29/100k | 2.7:1 | 0.000005 | corridor passenger connecting passengers corridors |
| 474.13/100k | 184.29/100k | 2.6:1 | 0.000000 | theory example currency absolute universally |
| 98.94/100k | 38.57/100k | 2.6:1 | 0.001628 | dell k rx sm borrow |
| 578.16/100k | 227.15/100k | 2.5:1 | 0.000000 | china ra ri kremlin european |
| 97.34/100k | 38.57/100k | 2.5:1 | 0.002130 | reporter rumour bush fox paul's |
| 139.93/100k | 55.72/100k | 2.5:1 | 0.000152 | defensive offense fighter target forces |
| 161.36/100k | 64.29/100k | 2.5:1 | 0.000056 | hero battlefield twilight doom quest |
| 520.22/100k | 210.00/100k | 2.5:1 | 0.000000 | effective result employ cumulative ease |
| 461.23/100k | 188.57/100k | 2.4:1 | 0.000000 | specialist financial service listed conceded |
| 226.32/100k | 94.29/100k | 2.4:1 | 0.000002 | role studied assistant former hopkins |
| 218.85/100k | 94.29/100k | 2.3:1 | 0.000009 | successful significance scholar considered accomplished |
| 382.38/100k | 167.15/100k | 2.3:1 | 0.000000 | counseling trusted desire statements religion |
| 104.42/100k | 47.14/100k | 2.2:1 | 0.004272 | losing future eager brink meantime |
| 245.78/100k | 111.43/100k | 2.2:1 | 0.000006 | shores wooded concrete slopes gliding |
| 222.87/100k | 102.86/100k | 2.2:1 | 0.000021 | positively psychological positive stress illness |
| 222.50/100k | 102.86/100k | 2.2:1 | 0.000027 | lore millions country cult decades |
| 174.40/100k | 81.43/100k | 2.1:1 | 0.000280 | electricity tires blade hammer panels |
| 255.01/100k | 120.00/100k | 2.1:1 | 0.000009 | wing slate split forming banners |
| 190.86/100k | 90.00/100k | 2.1:1 | 0.000123 | persisted reported preceded previous apparent |
| 118.04/100k | 55.72/100k | 2.1:1 | 0.003025 | squared vertical edges debris waves |
| 152.91/100k | 72.86/100k | 2.1:1 | 0.000764 | engagement demanded despised protested allowed |
| 151.86/100k | 72.86/100k | 2.1:1 | 0.000974 | specially recommended planned offer prepared |
| 332.41/100k | 162.86/100k | 2.0:1 | 0.000001 | harmless easier impossible frowned appropriately |
| 235.86/100k | 115.72/100k | 2.0:1 | 0.000047 | bacon cooking ate wholesome rice |
| 2225.32/100k | 1110.02/100k | 2.0:1 | 0.000000 | types such many other sometimes |
| 281.88/100k | 141.43/100k | 2.0:1 | 0.000011 | admirable amusing respond smarter troll |
| 220.09/100k | 111.43/100k | 2.0:1 | 0.000150 | errors error typed task tasks |
| 337.88/100k | 171.43/100k | 2.0:1 | 0.000002 | clicked note interrupt skipped missed |
| 142.57/100k | 72.86/100k | 2.0:1 | 0.003024 | mostly fairly slightly extraordinarily insanely |
| 150.85/100k | 77.14/100k | 2.0:1 | 0.002268 | ruthless elders captive disgrace slaves |
| 266.64/100k | 137.14/100k | 1.9:1 | 0.000034 | therefore nature ordinarily se leibniz |
| 324.42/100k | 180.00/100k | 1.8:1 | 0.000032 | needed raised flatland reared frail |
| 175.05/100k | 98.57/100k | 1.8:1 | 0.003617 | visitors midst neighbourhood photos barren |
| 1170.67/100k | 677.15/100k | 1.7:1 | 0.000000 | four second seventh s october |
| 276.55/100k | 167.15/100k | 1.7:1 | 0.000915 | pursue destined rewarded encourage foremost |
| 261.17/100k | 158.57/100k | 1.6:1 | 0.001301 | letter readers address pages link |
| 362.76/100k | 231.43/100k | 1.6:1 | 0.000478 | music onstage singer release violin |
| 494.22/100k | 321.43/100k | 1.5:1 | 0.000085 | passport paperwork ups debt credit |
| 991.36/100k | 677.15/100k | 1.5:1 | 0.000000 | wrong possible possibly agree sense |
| 34285.9/100k | 28898.9/100k | 1.2:1 | 0.000000 | brought had has for when |

## Over-sampled categories

| Corpus freq. | Stim. freq. | Ratio | p-value | Sample words |
|---|---|---|---|---|
| 127.32/100k | 1860.03/100k | 14.6:1 | 0.000000 | doin lookin hah suck dunno |
| 73.82/100k | 415.72/100k | 5.6:1 | 0.000000 | sixteen twenty-three eighteen seventy fifteen |
| 79.53/100k | 325.72/100k | 4.1:1 | 0.000009 | blouse skirts hat t-shirts apron |
| 27.27/100k | 98.57/100k | 3.6:1 | 0.000001 | sanchez merrill hal wharton neville |
| 116.92/100k | 381.43/100k | 3.3:1 | 0.000000 | hanging railing crawl doors chair |
| 34.25/100k | 107.14/100k | 3.1:1 | 0.003617 | gabrielle alice jane naomi bianca |
| 53.46/100k | 162.86/100k | 3.0:1 | 0.000190 | laundry upstairs bedroom drawers carpet |
| 25.76/100k | 72.86/100k | 2.8:1 | 0.000000 | nigh wasn idly heartily emphatically |
| 48.18/100k | 132.86/100k | 2.8:1 | 0.000764 | arrow claws pits dart claw |
| 263.90/100k | 608.58/100k | 2.3:1 | 0.000915 | screaming listen scream shouted wink |
| 332.17/100k | 732.87/100k | 2.2:1 | 0.000928 | friend's hers kids relax tomorrow |
| 144.10/100k | 291.43/100k | 2.0:1 | 0.002268 | beauty men fro spectacles boy's |
| 143.49/100k | 278.58/100k | 1.9:1 | 0.000000 | tired hurt tempted remembering remind |
| 126.86/100k | 240.00/100k | 1.9:1 | 0.000000 | foul herb cigarettes wash smelled |
| 251.57/100k | 467.15/100k | 1.9:1 | 0.000003 | son son's funeral grace death |
| 1994.22/100k | 3612.91/100k | 1.8:1 | 0.000000 | i'd i'm feeling absolutely sigh |
| 134.98/100k | 231.43/100k | 1.7:1 | 0.000280 | ronnie nigel dave bobby ted |
| 223.16/100k | 381.43/100k | 1.7:1 | 0.000047 | waits open opens motioned gestured |
| 269.01/100k | 428.58/100k | 1.6:1 | 0.000000 | exclaimed asked call wondering imagined |
| 322.88/100k | 510.01/100k | 1.6:1 | 0.000000 | hold blew rolled smashed roll |
| 2170.54/100k | 3390.05/100k | 1.6:1 | 0.000000 | until came old time |
| 173.72/100k | 270.00/100k | 1.6:1 | 0.000000 | dreamer best marvelous favorite marvellous |
| 4828.25/100k | 7418.68/100k | 1.5:1 | 0.000009 | want how saying things people |
| 205.97/100k | 312.86/100k | 1.5:1 | 0.000085 | fingernails shaving fingers mouth eyelids |
| 18745.95/100k | 27128.9/100k | 1.4:1 | 0.000000 | here keeps than much right |
| 607.73/100k | 865.73/100k | 1.4:1 | 0.000000 | knows nice pity wish happy |

**2. All words in each semantic cluster.** To interpret the semantic dimensions found using PCA, we clustered the words that had large projections on these dimensions into twelve clusters using *k*-means. This table lists all the words in each cluster and the label that was assigned to that cluster by the authors. Alternative methods for labeling the clusters are discussed and evaluated in Supplementary Data 5.

| Cluster Label | Cluster Words |
|---|---|
| *temporal* | travel minute leave date clock hours week rumbling next schedule months month immediately heading waited weeks weekend arrive seconds starting destination hour minutes twice parking trip halfway nights pm promptly |
| *abstract* | natural roots delicate exaggerated diverse gentle stronger atmosphere flesh soothing qualities muscular distorted describe strong powerful artificial deeper ecology stem pure sound spreading deep particularly intricate subtle masses expressive weak focus environment influenced hip creating intense sensation folk surroundings |
| *professional* | meetings owner worker office rented year business meet home decided visit staying paid bank students house members visiting meeting private staff school estate classroom college apartment hotel attend |
| *visual* | yellow fur silver badge garment large gold suit steel colour variety brown uniform cap clothing leather breeches coloured colored skull cotton wig bone wears fielder ribbons skin green stockings black seal breast glove stripes striped feathers jackets colors shafts white wide medium color style blazer shaped cloth tan |
| *violent* | lethal instantly breath kill bat painful pause repeat tongue stab trigger sentence breathe die accidentally poison accidental swallow explode bullet reaction kills hit swallowed repeatedly loses |
| *tactile* | fingers blade metallic fog melt slow vertical dome edges waves drifting absorb barrel inches flowing thin swirling smooth pinch diameter tops sliding thick gravity breeze depth drops lightly blades hits slowly surface thinner sheets heavier portable pressing needle solid cut thicker soft slight finger melting meters cloudy slowed lighter flow faster layers screens lighting inch clouds slower reach shapes stream layer upwards |
| *communal* | schools male community church young society interests bred banker family respected american culture catholic adopted whose teaching african among educated founded children public youth politician protect reputation sons wealthy |
| *mental* | asleep knew memories overwhelmed awake anxious uneasy studying moments hadn't learning sadness talked experiences sounded confess senses calm fascinating thoughts answered emptiness reading wake dreaming listening tense hearing experience awe reply exploring replies quiet solitude comforting wished explore happened realised discuss |
| *numeric* | four quarter pairs set pound pair five maximum extra half drop card pounds cent overhead deck floors three two per stock each top tie shillings purse ten twenty double sold intervals tables smaller six |
| *emotional* | alive nature innocent despised disgrace religion spiritual believes troubled emotion illness influence truth tortured compassion fear deeply speak perceived anger embrace religious emotions weakness human feelings harsh vile profoundly openly remark profound evil admiration believing peaceful christian convinced hatred man's cruel fearful betrayed |
| *social* | child son situation pleaded marriage parents arrest daughter victim husband informed charges charged suicide relatives sheriff widow accused met arrested confronted eldest father custody robbery pregnant murder mother guilty confessed calls wife court whom stolen refused married murderer murdered convicted |
| *locational* | stadium visitors halls company shops golf scenery architecture rooms gardens athletic lounge spacious space evenings building houses art landscape fields sporting shopping arts purchased annual national center stores pools campus facilities sports activities design uniforms clubhouse teams local |

**3. MNI coordinates for each identified semantic area.** The PrAGMATiC analysis found 77 significant semantic areas in the left hemisphere and 63 in the right. Each area was identified in the native anatomical space for each subject. This table lists the average MNI coordinates and the standard deviation for each area across subjects.

| Area | X (± std) | Y (± std) | Z (± std) |
|---|---|---|---|
| LPC L6 | -43±3.5 | -67±4.1 | 24±4.8 |
| LPC L9 | -51±5.1 | -57±4.4 | 23±5.2 |
| LPC L7 | -47±3.5 | -63±3.7 | 34±3.9 |
| LPC L11 | -53±1.8 | -54±3.7 | 36±6.3 |
| LPC L12 | -54±4.9 | -48±5.4 | 30±4.6 |
| LPC L3 | -36±5.9 | -78±4.9 | 32±4.3 |
| LPC L13 | -52±2.9 | -49±4.7 | 45±4.3 |
| LPC L14 | -57±1.9 | -41±2.7 | 31±4.6 |
| LPC L8 | -31±2.5 | -53±5.2 | 49±5.9 |
| LPC L5 | -31±3.1 | -65±3.2 | 45±4.2 |
| LPC L4 | -30±2.4 | -75±4.6 | 37±4.3 |
| LPC L1 | -37±3.2 | -81±2.6 | 21±4.1 |
| LPC L15 | -51±4.1 | -38±3.3 | 42±4.2 |
| LPC L10 | -38±5.0 | -50±5.9 | 44±3.9 |
| LPC L2 | -25±3.0 | -72±5.5 | 34±3.7 |
| MPC L10 | -5±1.5 | -55±3.2 | 26±2.9 |
| MPC L8 | -9±2.8 | -47±2.7 | 34±3.9 |
| MPC L6 | -5±1.1 | -59±1.6 | 37±4.3 |
| MPC L7 | -10±1.4 | -69±4.2 | 31±4.1 |
| MPC L5 | -8±2.1 | -49±3.4 | 44±3.0 |
| MPC L11 | -9±3.1 | -62±2.4 | 20±3.3 |
| MPC L14 | -6±3.8 | -45±3.3 | 9±5.3 |
| MPC L13 | -16±2.0 | -60±2.8 | 5±3.2 |
| MPC L12 | -8±2.7 | -55±2.2 | 12±3.7 |
| MPC L9 | -4±1.1 | -36±3.8 | 37±2.7 |
| MPC L1 | -8±2.4 | -57±4.6 | 58±3.0 |
| MPC L4 | -11±2.7 | -77±3.4 | 44±3.6 |
| MPC L3 | -5±2.5 | -66±2.8 | 49±4.2 |
| MPC L2 | -16±2.6 | -71±3.2 | 52±3.0 |
| SPFC L13 | -17±2.3 | 53±6.3 | 29±6.9 |
| SPFC L6 | -37±3.3 | 20±5.0 | 42±3.9 |
| SPFC L12 | -5±0.8 | 54±4.7 | 33±4.5 |
| SPFC L15 | -6±1.1 | 56±6.1 | 18±3.2 |
| SPFC L4 | -34±5.0 | 19±5.4 | 48±6.9 |
| SPFC L10 | -6±1.5 | 40±3.8 | 46±5.5 |
| SPFC L5 | -9±2.4 | 26±3.4 | 57±2.8 |
| SPFC L14 | -8±2.6 | 47±4.1 | 18±7.5 |
| SPFC L1 | -8±2.9 | 10±2.1 | 65±1.5 |
| SPFC L18 | -6±1.8 | 60±2.3 | -8±3.3 |
| SPFC L17 | -7±1.3 | 51±4.4 | 0±2.7 |
| SPFC L11 | -23±2.7 | 43±3.9 | 35±3.6 |
| SPFC L9 | -16±3.6 | 37±3.3 | 45±4.1 |
| SPFC L16 | -15±3.4 | 67±2.8 | 11±5.2 |
| SPFC L7 | -22±2.0 | 23±2.5 | 46±3.6 |
| SPFC L2 | -19±2.4 | 12±3.9 | 61±4.0 |
| SPFC L8 | -33±1.7 | 32±3.6 | 35±4.3 |
| SPFC L3 | -27±1.9 | 8±4.4 | 52±1.4 |
| LTC L6 | -53±5.4 | -29±6.1 | -3±2.4 |
| LTC L8 | -51±3.6 | 3±6.6 | -21±6.0 |
| LTC L4 | -56±3.2 | -45±4.9 | 3±3.8 |
| LTC L7 | -59±2.5 | -8±6.1 | -23±6.4 |
| LTC L5 | -61±3.5 | -35±7.0 | -7±2.9 |
| LTC L3 | -57±3.3 | -56±6.3 | -1±4.1 |
| LTC L1 | -48±4.4 | -64±4.8 | 1±4.8 |
| LTC L2 | -54±4.8 | -59±5.2 | -9±4.4 |
| VTC L6 | -33±4.1 | -23±1.6 | -22±1.3 |
| VTC L5 | -29±1.1 | -39±1.5 | -13±0.9 |
| VTC L1 | -47±5.1 | -66±4.1 | -9±5.3 |
| VTC L4 | -43±2.1 | -48±4.6 | -14±2.6 |
| VTC L3 | -50±3.6 | -55±5.4 | -14±4.9 |
| VTC L2 | -45±3.9 | -59±4.3 | -11±2.4 |
| IPFC L9 | -34±5.0 | 57±5.6 | 3±6.5 |
| IPFC L8 | -45±2.9 | 37±2.2 | -7±2.3 |
| IPFC L11 | -38±2.0 | 53±2.2 | -10±3.1 |
| IPFC L12 | -22±3.0 | 51±2.9 | -17±1.2 |
| IPFC L1 | -42±3.1 | 3±4.9 | 40±5.3 |
| IPFC L10 | -35±1.9 | 39±2.6 | -14±2.1 |
| IPFC L5 | -41±4.0 | 35±5.8 | 26±3.8 |
| IPFC L7 | -41±4.2 | 42±4.7 | 4±3.3 |
| IPFC L3 | -46±4.1 | 7±4.5 | 22±5.4 |
| IPFC L2 | -43±5.2 | 8±6.2 | 31±2.7 |
| IPFC L4 | -45±3.5 | 28±7.4 | 19±6.1 |
| IPFC L6 | -44±2.5 | 38±7.9 | 14±2.3 |
| OIC L1 | -42±1.8 | -15±4.1 | 47±4.8 |
| OIC L2 | -54±2.6 | 0±3.3 | 7±3.8 |
| OIC L3 | -35±1.5 | 14±5.4 | 6±3.1 |
| OIC L4 | -39±1.0 | -2±4.5 | -6±2.3 |
| LPC R7 | 52±4.7 | -53±3.6 | 28±4.7 |
| LPC R5 | 48±5.3 | -58±3.7 | 25±4.8 |
| LPC R8 | 43±4.1 | -61±4.2 | 45±4.2 |
| LPC R9 | 50±5.4 | -52±3.7 | 42±5.1 |
| LPC R10 | 54±5.3 | -46±2.8 | 32±2.8 |
| LPC R3 | 45±4.0 | -68±2.7 | 26±4.9 |
| LPC R2 | 50±2.8 | -61±4.1 | 7±5.9 |
| LPC R12 | 55±4.7 | -41±3.2 | 42±4.9 |
| LPC R4 | 36±3.7 | -70±4.2 | 39±4.3 |
| LPC R1 | 41±3.9 | -76±3.8 | 20±2.9 |
| LPC R13 | 55±2.6 | -33±5.2 | 44±4.5 |
| LPC R11 | 41±5.5 | -45±2.8 | 42±3.7 |
| LPC R6 | 31±3.9 | -64±3.4 | 44±5.9 |
| MPC R7 | 9±1.0 | -52±3.7 | 33±2.8 |
| MPC R6 | 5±0.8 | -60±3.5 | 27±2.5 |
| MPC R3 | 12±1.8 | -69±4.9 | 36±5.9 |
| MPC R10 | 4±0.2 | -29±2.3 | 26±1.2 |
| MPC R8 | 11±1.5 | -57±2.4 | 17±3.3 |
| MPC R4 | 7±0.8 | -55±3.2 | 45±2.8 |
| MPC R5 | 10±3.0 | -45±4.7 | 54±5.4 |
| MPC R2 | 11±1.3 | -58±4.0 | 61±2.6 |
| MPC R9 | 6±1.1 | -37±2.5 | 36±2.4 |
| MPC R1 | 21±3.4 | -74±5.5 | 41±2.0 |
| SPFC R9 | 38±4.1 | 16±6.8 | 38±3.1 |
| SPFC R11 | 37±2.8 | 27±5.1 | 36±4.4 |
| SPFC R17 | 28±1.5 | 58±2.4 | 7±4.9 |
| SPFC R15 | 20±2.0 | 52±2.8 | 31±3.5 |
| SPFC R8 | 12±3.3 | 25±3.7 | 57±3.2 |
| SPFC R16 | 9±1.2 | 57±2.7 | 17±1.7 |
| SPFC R14 | 7±0.7 | 45±4.4 | 40±2.4 |
| SPFC R12 | 17±2.5 | 40±5.1 | 47±2.8 |
| SPFC R6 | 39±2.9 | 18±5.8 | 48±5.1 |
| SPFC R19 | 9±2.4 | 66±2.9 | -6±5.6 |
| SPFC R2 | 10±3.1 | 8±4.8 | 65±4.4 |
| SPFC R18 | 8±0.6 | 48±2.0 | -4±1.8 |
| SPFC R4 | 41±1.9 | 2±4.4 | 46±6.2 |
| SPFC R10 | 24±3.5 | 29±3.4 | 44±4.9 |
| SPFC R3 | 19±3.0 | 14±4.9 | 62±3.3 |
| SPFC R5 | 33±3.3 | 8±5.5 | 52±2.7 |
| SPFC R1 | 23±1.3 | 3±4.6 | 59±3.4 |
| SPFC R13 | 28±4.1 | 35±3.9 | 38±5.0 |
| SPFC R7 | 23±2.6 | 18±4.0 | 50±3.6 |
| LTC R5 | 57±3.9 | -23±4.9 | -6±3.8 |
| LTC R7 | 51±2.3 | 11±6.1 | -21±4.8 |
| LTC R4 | 53±5.1 | -27±6.4 | 1±3.7 |
| LTC R3 | 57±3.5 | -41±4.3 | 5±4.0 |
| LTC R8 | 47±1.9 | 7±4.0 | -35±2.1 |
| LTC R6 | 63±4.4 | -25±5.3 | -13±4.4 |
| LTC R2 | 57±3.1 | -51±6.0 | -1±4.4 |
| LTC R1 | 55±3.7 | -57±5.7 | -6±4.5 |
| VTC R2 | 42±1.2 | -14±5.1 | -31±2.5 |
| VTC R1 | 30±1.2 | -25±3.7 | -21±2.0 |
| IPFC R7 | 49±2.4 | 35±3.2 | -1±2.2 |
| IPFC R4 | 52±3.2 | 25±3.2 | 13±3.0 |
| IPFC R8 | 41±2.3 | 54±2.3 | -3±4.1 |
| IPFC R3 | 46±2.5 | 29±3.5 | 28±4.0 |
| IPFC R9 | 31±1.9 | 41±3.3 | -14±1.5 |
| IPFC R2 | 51±3.2 | 11±3.3 | 16±2.7 |
| IPFC R6 | 43±3.1 | 46±2.8 | 11±5.4 |
| IPFC R1 | 42±3.6 | 11±3.1 | 31±2.6 |
| IPFC R5 | 45±4.0 | 36±4.5 | 15±3.6 |
| OIC R3 | 55±4.9 | 4±3.7 | 16±5.4 |
| OIC R2 | 38±0.9 | 15±4.1 | 5±1.7 |
| OIC R1 | 40±0.4 | -3±3.1 | -3±2.3 |