Title:
A functional dissociation between language and multiple-demand systems revealed in patterns of BOLD signal fluctuations

Abbreviated title:
Language and MD systems functionally dissociated

Authors and affiliation:
Idan Blank, Nancy Kanwisher and Evelina Fedorenko
Brain and Cognitive Sciences Department and McGovern Institute of Brain Research, Massachusetts Institute of Technology, 43 Vassar Street, Cambridge, MA 02139, USA

Corresponding Author:
Idan Blank, Department of Brain and Cognitive Sciences, MIT, 43 Vassar Street, 46-4141f, Cambridge, MA 02139, USA, iblank@mit.edu

Author contributions:
I.B. analyzed the data; N.K. supervised the project; E.F. developed tasks and stimuli, collected fMRI data and designed the procedure for defining functional regions of interest. All authors wrote the manuscript.

51 pages | 2 tables | 6 figures

Abstract: 186 words
Abstract

What is the relationship between language and other high-level cognitive functions? Neuroimaging studies have begun to illuminate this question, revealing that some brain regions are quite selectively engaged during language processing whereas other, "multiple-demand (MD)" regions are broadly engaged by diverse cognitive tasks. Nonetheless, the functional dissociation between the language and MD systems remains controversial. Here we tackle this question with a synergistic combination of fMRI methods: we first define candidate language-specific and MD regions in each subject individually (using functional localizers), and then measure BOLD signal fluctuations in these regions during two naturalistic conditions (“rest” and story-comprehension). In both conditions, signal fluctuations strongly correlate among language regions, as well as among MD regions, but correlations across systems are weak or negative. Moreover, data-driven clustering analyses based on these inter-region correlations consistently recover two clusters corresponding to the language and MD systems. Thus, although each system forms an internally integrated whole, the two systems dissociate sharply from each other. This independent recruitment of the language and MD systems during cognitive processing is consistent with the hypothesis that these two systems support functionally distinct cognitive functions.

Keywords: Functional connectivity, language, multiple demand system
Introduction

Although the key brain regions engaged in language processing have been known since Broca (1861/2006) and Wernicke (1874/1969), debates continue on whether, and to what extent, they overlap with regions engaged in other cognitive processes. Many neuroimaging studies have reported that brain regions that process language (particularly Broca’s area) also engage in many non-linguistic processes, including arithmetic (e.g., Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999; Stanescu-Cosson et al., 2000), music perception (Koelsch, Gunter, Zysset, Lohmann, & Friederici, 2002; Maess, Koelsch, Gunter, & Friederici, 2001), working memory and cognitive control (e.g., Blumstein, 2009; Hein & Knight, 2008; January, Trueswell, & Thompson-Schill, 2009; Kaan & Swaab, 2002; Koechlin & Jubault, 2006). Yet other studies have found that regions activated during non-linguistic tasks are distinct from language-processing regions (Fedorenko, Behr, & Kanwisher, 2011; Fedorenko, Duncan, & Kanwisher, 2012; Monti & Osherson, 2012; Monti, Parsons, & Osherson, 2009, 2012). Specifically, many cognitively demanding tasks activate a set of frontal and parietal regions known as the "multiple demand (MD)" system (Braver, Reynolds, & Donaldson, 2003; Cole & Schneider, 2007; Dosenbach, Fair, Cohen, Schlaggar, & Petersen, 2008; Duncan, 2010; Duncan & Owen, 2000; Fedorenko, Duncan, & Kanwisher, 2013; Miller & Cohen, 2001), which does not overlap with the classic fronto-temporal language system. Nonetheless, the dissociation between a putatively language-specific system and this domain-general MD system remains controversial (Blumstein & Amso, 2013; Thompson-Schill, Bedny, & Goldberg, 2005).
To test for this dissociation, here we compared the BOLD signal time-courses of candidate language and MD regions by synergistically combining two fMRI methods: functional localizers and functional correlations. First, we functionally localized candidate regions of interest in each subject (Fedorenko et al., 2013; Fedorenko, Hsieh, Nieto-Castañón, Whitfield-Gabrieli, & Kanwisher, 2010) using tasks that target linguistic processing (language localizer) and cognitive effort (MD localizer). Next, subjects were scanned during a "rest" period or during a story-comprehension task (conditions that are independent of, and less constrained than, the localizers). The time-course of BOLD signal fluctuations during these two conditions was then extracted from each functionally-defined region. Finally, we measured the pairwise correlations between time-courses of different regions (separately for each condition). This approach enabled us to answer three questions: (i) To what extent do candidate language regions form a functionally integrated system (Cordes et al., 2000; Hampson, Peterson, Skudlarski, Gatenby, & Gore, 2002; Newman, Kenny, Saint-Aubin, & Klein, 2013; Turken & Dronkers, 2011; Yue, Zhang, Xu, Shu, & Li, 2013), as indexed by high correlations among these regions?; (ii) To what extent do candidate MD regions similarly form an integrated system? (Dosenbach et al., 2007; Hampshire, Highfield, Parkin, & Owen, 2012; Seeley et al., 2007); and critically, (iii) How functionally dissociable are language and MD regions from each other, as indexed by weak (or negative) correlations between pairs of regions straddling the two systems?

The current approach harnesses the complementary strengths of functional localizers and functional correlations. First, the “rest” and story-comprehension conditions allow us to sample a broader, more naturalistic range of cognitive
processes compared to task-based studies. Second, functional correlations allow us to use not only hypothesis-driven methods but also data-driven clustering to discover the relationship between language and MD regions based on the co-variation of their respective signal time-courses. Third, we can straightforwardly interpret the emerging clusters in terms of their functionally characterized constituents, because our regions of interest are functionally localized. Therefore, we do not have to rely on "reverse inference" from stereotaxic coordinates (Poldrack, 2006), which is inevitable when no functional localizers are used (Lee et al., 2012; Mantini, Corbetta, Romani, Orban, & Vanduffel, 2013; Tie et al., 2012; Yeo et al., 2011).

Materials and methods

Subjects

Eighteen adult subjects (6 males) aged 18-30 participated in a resting-state scan. Ten subjects (3 males) aged 18-30 were scanned during a story-comprehension task (six of these also participated in the resting-state scan). All 22 subjects also completed independent localizer runs that were used to define candidate language and MD regions. Subjects were right-handed, native speakers of English from MIT and the surrounding Cambridge community. All provided informed consent in accordance with the requirement of the Internal Review Board at MIT, and were paid for participating.

Functional localization of candidate language and MD regions
Data acquisition. Structural and functional data were collected on a whole-body 3 Tesla Siemens Trio scanner with a 32-channel head coil at the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research at MIT. T1-weighted structural images were collected in 176 sagittal slices (1mm isotropic voxels; TR = 2530ms; TE = 3.48ms). Functional blood oxygenation level dependent (BOLD) data were acquired using an EPI sequence with a flip angle of 90° and applying GeneRalized Autocalibrating Partially Parallel Acquisition (GRAPPA) with an acceleration factor of 2. Images were collected in 31 near-axial slices, acquired in an interleaved order with a 10% distance factor (in-plane resolution: 2.1×2.1mm; slice thickness: 4mm; FoV: 200mm in the phase encoding (A>>P) direction; matrix size: 96mm×96mm; TR: 2000ms; TE: 30ms). Prospective Acquisition CorrEction (PACE; Thesen, Heid, Mueller, & Schad, 2000) was used to adjust the positions of the gradients based on the subject’s head motion one TR back. The first ten seconds of each run were excluded to allow for steady state magnetization.

Design. Candidate language and MD regions of interest were functionally defined. The language localizer has been previously demonstrated to identify language-selective brain regions (Fedorenko et al., 2011) sensitive to high-level linguistic information, including syntax and lexical semantics (Fedorenko, Nieto-Castanon, & Kanwisher, 2012) and, accordingly, contrasted reading of sentences with reading of sequences of pronounceable nonwords (presented one word/nonword at a time). After each sentence/nonword sequence, a probe word/nonword appeared and subjects had to decide whether the probe item appeared in the preceding stimulus. Each of our 22 subjects completed 2-4 runs of the localizer, with Sentence and Nonwords blocks lasting either 24s (4 blocks per
condition per run in a 336s run, 4 subjects) or 18s (8 blocks per condition per run in a 378s run, 12 subjects; or 6 blocks per condition per run in a 396s run, 6 subjects). The order of Sentence and Nonwords blocks was counterbalanced across runs and subjects. This localizer (available from http://web.mit.edu/evelina9/www/funcloc/funcloc_localizers.html) is robust to changes in materials, task and modality of presentation (Fedorenko et al., 2010).

For the MD localizer, subjects performed a spatial working memory task that we have found to broadly and robustly activate the MD system (Fedorenko et al., 2013). Subjects had to keep track of four (easy condition) or eight (hard condition) locations in a 3×4 grid (Fedorenko et al., 2011). In both conditions, subjects performed a two-alternative forced-choice task at the end of each trial to indicate the set of locations they just saw. The contrast Hard > Easy targets brain regions engaged in cognitively demanding tasks; Fedorenko et al. (2013) have shown that the regions activated by this task are also activated by a wide range of other tasks contrasting a difficult versus an easy condition. Each of our 22 subjects completed 1-3 runs of this MD localizer, with Hard and Easy blocks lasting either 34s (5 blocks per condition per run in a 436s run, 10 subjects), 32s (6 blocks per condition per run in a 448s run, 11 subjects) or 18s (6 blocks per condition per run in a 288s run, 1 subject). The order of hard and easy blocks was counterbalanced across runs and subjects.

**Spatial image preprocessing and localizer task analysis.** MRI data were spatially preprocessed using SPM5 (http://www.fil.ion.ucl.ac.uk/spm). To reduce data transformations, we performed all analyses in native functional space. Each subject’s functional data were motion-corrected and then smoothed with a 4mm
FWHM Gaussian filter. In addition, the anatomical image of each subject was segmented into three probability maps denoting areas of gray matter, white matter and CSF (see functional correlation analysis), and these maps were then co-registered to the native functional space. The anatomical images were also further used for cortical surface reconstruction (see next section). The data for the localizer tasks were modeled using a general linear model with a boxcar regressor convolved with a canonical HRF function.

**Definition of group-constrained, subject-specific fROIs.** For each subject, functional regions of interest (fROIs) were defined by combining two sources of information (Fedorenko et al., 2010; Julian, Fedorenko, Webster, & Kanwisher, 2012): (i) the subject’s own activation map from the localizer runs; and (ii) group-level constraints (“masks”). The latter demarcated brain areas within which most or all individuals in prior studies showed activity for the localizer contrasts (Table 1).

For the language fROIs, we used masks derived from a group-level representation of data for the Sentences > Nonwords contrast in an independent group of subjects (Fedorenko et al., 2010; masks available for download at http://web.mit.edu/evelina9/www/funcloc/funcloc_parcels.html). Following Fedorenko et al. (2011), eight masks were used for defining fROIs in the left hemisphere, including regions in the posterior and middle-posterior temporal lobe (PostTemp and MidPostTemp, respectively), anterior and middle-anterior temporal lobe (AntTemp and MidAntTemp, respectively), inferior and orbital-inferior frontal gyrus (IFG and IFGorb, respectively), middle frontal gyrus (MFG) and angular gyrus (AngG). These masks were mirror-projected onto the right hemisphere to create eight homologous masks, resulting in a total of 16 language masks. As the masks
(denoted by gray contours in Figure 1) cover significant parts of the cortex, their mirrored version is likely to encompass the right hemisphere homologues of the left-hemispheric language-selective regions, despite possible hemispheric asymmetries in their precise anatomical location.

The right-hemisphere homologues were included for two reasons. First, these regions appear to be activated during at least some aspects of language processing, albeit usually not as strongly as the typical, left-lateralized language regions (Chiarello, Banich, & Mack, 2003; Jung-Beeman, 2005). Second, given that left hemisphere damage, but typically not right hemisphere damage, leads to difficulties in language production and comprehension (Damasio, 1992; Geschwind, 1970), we wanted to examine hemispheric differences in functional correlations. For example, we wanted to see whether left hemisphere language regions would show stronger inter-region correlations, which might be expected of a core language system (cf. the MD regions, which are strongly bilateral and should thus show similarly strong inter-region correlations in the left and right hemispheres).

For the MD fROIs, we used anatomical masks (Tzourio-Mazoyer et al., 2002) that included the main regions linked to MD activity in prior work, following the methods of Fedorenko et al. (2013) (for a similar approach, see: Fedorenko, Duncan, et al., 2012). Eighteen masks were used, 9 in each hemisphere, including regions in the opercular inferior frontal gyrus (IFGop), middle frontal gyrus including its orbital part (MFG and MFGorb, respectively), insular cortex (Insula), precentral gyrus (PrecG), supplementary and pre-supplementary motor area (SMA), inferior and superior parietal lobe (ParInf and ParSup, respectively) and anterior cingulate cortex (ACC).
These group-level masks, in the form of binary maps, were used to systematically constrain the selection of subject-specific fROIs. Thus, for each subject, 16 candidate language fROIs were created by intersecting the subject's unthresholded \( t \) map for the Sentences > Nonwords contrast with the language masks. For each subject and for each mask, the 10% of voxels with the highest \( t \)-values in the intersection image were then used to define a fROI (note that the voxels included in the right hemisphere fROIs were not constrained to be mirrored versions of their left hemisphere counterparts, but were only constrained to land within a mirrored version of the broad masks). Similarly, 18 candidate MD fROIs were created for each subject by intersecting the subject's unthresholded \( t \) map for the Hard > Easy contrast with the MD masks, again selecting the 10% of voxels with the highest \( t \)-values within each mask. Finally, we excluded a small set of voxels that were contained in more than one fROI due to small spatial overlap between language and MD activation maps. Across subjects, these excluded voxels comprised 1.67% (2.60) of our fROIs (for all subjects, 16 fROIs had no voxels excluded from them; each of the remaining 18 fROIs had less than 7% excluded voxels).

Defining fROIs by choosing the 10% of voxels with the highest \( t \)-values for a localizer contrast in a given mask balances the trade-off between: (i) choosing only voxels with a BOLD time-course that strongly co-varies with the localizer conditions (Sentences > Non-words or Hard > Easy); and (ii) having a sufficient number of voxels in the fROI. In addition, this procedure ensures that each fROI has a constant size across subjects. However, we obtained similar results to those reported below when fROIs were instead defined by intersecting the language or MD masks with a
thresholded $t$-map for the language or MD contrasts, respectively ($p = 0.001$, uncorrected).

The language and MD masks used here were originally created in Montreal Neurological Institute (MNI) space. Therefore, prior to defining fROIs, the masks had to be projected onto each subject’s native functional space. This was done in two steps: first, Combined Volume and Surface (CVS) registration (Postelnicu, Zollei, & Fischl, 2009) was used to estimate the transformation of an MNI anatomical (T1) template to the native anatomical space of each subject, and the resulting transformation was applied to the masks. Second, affine co-registration was used to project the masks from native anatomical space onto native functional space. Only then were the masks intersected with subjects’ $t$ maps from the functional localizers.

The localizer effects were highly reliable in all fROIs. Reliability was tested via an $n$-fold, leave-one-out cross validation across runs: for each subject, we defined fROIs based on all localizer runs but one, and then derived estimates of the localizer contrast effect for the left-out run in these fROIs. The contrast effect estimates were averaged across all possible left-out runs and then tested for significance across subjects (FDR-corrected for the number of regions). For all left-hemisphere candidate language regions, $t_{(21)} > 6.47, p < 10^{-5}$. For all right-hemisphere candidate language regions, $t_{(21)} > 2.70, p < 0.007$. For all candidate MD fROIs, $t_{(19)} > 5.64, p < 10^{-4}$. Figure 1 shows the language and MD fROIs in the left hemisphere of 3 representative subjects, as well as probability maps of fROI locations across all subjects.
Functional correlation analysis

Data acquisition. Functional data were collected using the same parameters as for the functional localizers.

Design. In the resting-state condition, subjects were instructed to close their eyes but to remain awake and let their mind wonder for 5mins. In the story-comprehension condition, subjects listened to 4-6 stories over the scanner-safe headphones (Sensimetrics). Each story lasted between 4.5 and 6mins. Stories were constructed from existing publicly available texts (fairy-tales, short stories and Wikipedia articles) but edited so as to include a variety of linguistically interesting phenomena that do not occur with sufficiently high frequency in natural texts (e.g., infrequent words, non-local syntactic dependencies, unusual syntactic constructions, temporary ambiguity, etc.; for examples, see Table 2). (The motivation for editing the stories had to do with the use of these materials in another project aimed at understanding the processing of different kinds of linguistic complexity.) The stories were then recorded by two native English speakers (one male, one female). After each story, subjects answered six comprehension questions, presented in a two-alternative forced-choice format. For each subject, accuracy on these questions was significantly above chance, as indicated by the binomial test (for all tests, \( p < 10^{-11} \), Bonferroni corrected for the number of subjects; mean accuracy across subjects: 83.09\%(10)).

Spatial image preprocessing. Functional data were spatially preprocessed using the same procedure applied for the localizer runs.
Temporal preprocessing. Temporal preprocessing was carried out using the CONN toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012) with default parameters unless otherwise specified. First, noise introduced by signal fluctuations originating from non-neuronal sources (e.g., respiratory and cardiac activity) was removed. To this end, the first 5 temporal principal components of the BOLD signal time-course extracted from the white matter were regressed out of each voxel’s time-course; signal originating in the CSF was similarly regressed out. White-matter and CSF voxels were identified based on segmentation of the anatomical image (Behzadi, Restom, Liau, & Liu, 2007). The first 6 principal components of the six motion parameters estimated during offline motion correction were also regressed out, as well as their first time-derivative. Second, the residual signal was bandpass filtered (0.008-0.09Hz) in order to only preserve low-frequency signal fluctuations (Cordes et al., 2001).

Data analysis. Analysis of functional correlations was carried out separately for each of the two experimental conditions (resting-state and story-comprehension). For each subject, we averaged the BOLD signal time-course across all voxels in each fROI. For each pair of fROIs, Pearson’s moment correlation coefficient was then computed between their respective time-courses. These correlations were Fisher-transformed to improve normality, and three one-way, repeated-measures Analyses of Variance (ANOVAs) were then performed on the data in order to assess whether different regions showed different patterns of functional correlations. First, we compared the average correlation within the candidate language system (i.e., the average of all 120 pairwise correlations among the 16 language fROIs), the average correlation within the candidate MD system (i.e.,
the average of all 153 pairwise correlations among the 18 MD fROIs), and the average correlation across the two systems (i.e., the average of all 288 pairwise correlations between a language fROI and a MD fROI). Second, for the language system, we compared the mean correlation within the left-hemisphere (averaging across all 28 pairwise correlations among the 8 language fROIs in the left-hemisphere), the analogous mean correlation within the right-hemisphere, and the mean correlation across the two hemispheres (averaging across 64 pairwise correlations). Third, the same within-hemisphere vs. across-hemispheres comparison was carried out for the MD system.

For visualization purposes, we also created two group-level matrices of fROI-to-fROI correlations, one for each experimental condition. Specifically, the Fisher-transformed correlation between each pair of fROIs was averaged across subjects (the Fisher-transform decreases the bias in averaging; Silver & Dunlap, 1987), and the resulting average correlations were then inverse-Fisher-transformed. The two group-level fROI-to-fROI correlation matrices are presented in Figure 2. However, to ensure that the patterns of functional correlations reported here were consistently observed across individual subjects, the majority of our analyses did not use these average correlation matrices but were instead performed within subjects (e.g., the repeated-measures ANOVAs described above). Only the hierarchical clustering analysis (see below) relied on group-level average correlations.

Controlling for the effects of head-motion. Head-motion artifacts have been reported to affect functional correlations (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012; Satterthwaite et al., 2012; Van Dijk et al., 2010). To ensure that such
artifacts could not account for the results reported here, we performed two control
analyses. First, for each experimental run of each subject, time-points with excessive
head-motion ("motion spikes") were identified using the Artifact Detection Toolbox
implemented in MATLAB (available for download at http://www.nitrc.org/projects/artifact_detect/). Each motion spike was then
included as a regressor during temporal preprocessing (see above), thus removing
the effects of these time-points on our residual BOLD time courses (Lemieux, Salek-
Haddadi, Lund, Laufs, & Carmichael, 2007; Satterthwaite et al., 2013). This control
analysis and our original analysis (without "spike regression") resulted in
qualitatively similar patterns of functional correlations. To minimize data
manipulations, we report the results of our original analysis.

Second, we tested whether individual differences in estimated head-motion
could explain individual differences in functional correlation patterns. First, each
subject’s six motion parameters were collapsed to a single value, Mean Relative
Displacement (MRD; Jenkinson, Bannister, Brady, & Smith, 2002). Then, we
computed the Pearson correlation, across subjects, between MRD and each of three
measures described above: (i) the mean functional correlation within the language
system; (ii) the mean functional correlation within the MD system; and (iii) the
mean functional correlation across the two systems. None of these measures were
significantly correlated with MRD (FDR-corrected) in either the resting-state or the
story-comprehension conditions (the correlations that were significant prior to FDR
correction were opposite in direction to our reported effects). On average,
individual differences in head-motion explained 5% of the individual differences in
functional correlation patterns. We then repeated this analysis using the point bi-
serial correlation instead of the Pearson correlation by splitting our sample into a
“high MRD” half and a “low MRD” half. While the two halves significantly differed in
MRD, they did not differ in their functional correlation patterns. We therefore
conclude that our results reported below cannot result from a head-motion artifact.

Clustering analyses

K-means. In order to reveal the dominant patterns of functional correlations across
our fROIs in a relatively data-driven fashion, we submitted the average BOLD signal
time-courses from candidate language and MD fROIs to Matlab’s K-means clustering
algorithm, separately for each subject. To ensure the choice of $K$ did not impose on
the data an implicitly hypothesized division into language and MD systems, separate
analyses were run with values of $K$ ranging from 2 to 6 clusters. The pairwise
distance measure used for clustering was defined as 1 minus the correlation
between different time-courses. To choose the initial cluster centroids, $K$ time-
courses out of a subject’s data were randomly sampled, and this procedure was
repeated 50 times to generate multiple clustering solutions. To pool the resulting
data we then computed, for each pair of fROIs, the probability (percentage of
solutions across random initializations and subjects) that the two fROIs would both
be assigned to the same cluster. Such pooling provides a straightforward way to
collapse results across subjects with different cluster solutions or cluster-
numbering (i.e., order) (such pooling is also known as “consensus clustering”, see:
Bassett et al., 2013; Lancichinetti & Fortunato, 2012).

We assessed the significance of our clustering results with a permutation
test. For each subject, we created a surrogate BOLD signal time-course for each fROI
by phase-shuffling its original time-course (i.e., reassigning the phases of different
frequencies uniformly at random, with replacement). The 34 surrogate time-courses
of each subject were then clustered and the clustering solutions were pooled across
subjects, using the same procedures described above. We repeated this clustering of
surrogate time-courses 1,000 times, generating for each pair of fROIs a null
distribution of the probability that they would be assigned to the same cluster. The
true probability, based on clustering the real data, was compared against this
distribution to produce a two-tailed $p$-value. Multiple comparisons were FDR-
corrected (Benjamini & Yekutieli, 2001) separately for each fROI (each pair of the 34
fROIs was assigned a $p$-value, so each fROI had 33 $p$-values associated with it). This
test yielded results similar to those obtained when we shuffled the original time-
courses across fROIs instead of generating surrogate time-courses.

In the previous analysis, the clustering algorithm was data-driven in the
sense that it was not provided with information about which fROIs were candidate
language regions and which were candidate MD regions. Nonetheless, the analysis
was still constrained to treat each fROI as a distinct entity, as the clustering was run
on time-courses that were averaged across all voxels within each fROI. To partially
relax this constraint, our next analysis clustered the BOLD signal time-courses of
individual voxels across all fROIs. As in the previous analysis, 50 clustering solutions
were generated for each subject, where randomly sampled time-courses served as
initial cluster centroids. For each clustering solution, we then performed the
following computation: first, we examined each set of voxels originating from within
a single fROI to determine its “dominant cluster” (that is, the cluster that had the
largest number of voxels in that fROI assigned to it). Then, for each pair of fROIs, we
computed the percentage of voxels in the first fROI assigned to the cluster that was
dominant in the second fROI (this resulted in two measures, depending on which
fROI was "first" and which was "second"). This procedure provided, for each pair of
fROIs, a voxel-wise measure of cluster similarity, which was then averaged across
random initializations and subjects. To test the significance of this cluster similarity
measure for each pair of fROIs, we applied a phase-shuffling permutation test
following the same procedure as described above.

This second analysis used data from single voxels rather than fROIs, yet the
similarity measure we computed for pooling the results across subjects still referred
to the original grouping of voxels into fROIs. However, this measure was only
computed after voxel-wise clustering had taken place. We chose this measure
because our data did not allow for comparing single voxels across subjects: as fROIs
were defined in a subject-specific manner, voxels falling within a fROI for one
subject might not have fallen within any fROI for another subject, thus entering the
clustering analysis for the former subject but not the latter.

Hierarchical clustering. Hierarchical clustering is an algorithm that creates a
binary tree-structure connecting elements in a set, such that the length of branches
on the tree approximates the distances among the elements as provided by the user
(Hartigan, 1975). Clustering together elements whose connecting path on the tree is
shorter than a chosen length therefore creates a partition of the element set without
pre-specifying the number of resulting clusters (in contrast to $K$-means). We
performed hierarchical clustering on our fROIs, providing the group-level fROI-to-
fROI correlation matrix as input, so that the distance between two fROIs was defined
as 1 minus their correlation. Clustering was based on average linkage, so that two
clusters were merged into a bigger cluster based on the mean distance between their respective members.

The optimal partition of fROIs based on the resulting tree was identified via a measure of modularity (Newman & Girvan, 2004). First, by gradually decreasing the path-length used as a criterion for clustering fROIs, we generated the set of all possible partitions licensed by our hierarchical clustering solution (the longest path-length generates a single cluster consisting of all 34 fROIs; the shortest path-length generates 34 singleton clusters). Then, for each partition, we computed a reformulated modularity measure that is appropriate for detecting clusters in correlated data (Gómez, Jensen, & Arenas, 2009). High modularity values indicate clustering solutions where, within each cluster, the positive functional correlations are stronger (and the negative functional correlations are weaker) compared to what is expected under a null model. The null model is a random fROI-to-fROI correlation matrix that preserves, for each fROI, the sum of its positive correlations and the sum of its negative correlations with the other fROIs.

Results

Functional correlation analysis: comparing systems and hemispheres

Comparison of the language and MD systems

Figure 2 presents, for the resting-state and story-comprehension conditions, matrices of pairwise correlations between candidate language and MD fROIs
computed on the time-courses of BOLD signal fluctuations. A clear partition of the
fROIs is visually evident prior to any statistical analysis: most pairs of language
fROIs are strongly and positively correlated with each other (cf. top-left quadrant of
the correlation matrices), and most pairs of MD ROIs are also strongly and positively
correlated with each other (cf. bottom-right quadrant of the correlation matrices);
but correlations of most pairs consisting of a language fROI and a MD fROI are
noticeably weaker (cf. bottom-left and top-right quadrants of the correlation
matrices). Moreover, during story-comprehension there are significant negative
correlations between candidate left-hemisphere language fROIs and right-
hemisphere MD fROIs. These results indicate a functional architecture comprised of
two systems, one consisting of language regions and the other consisting of MD
regions. BOLD signal fluctuations within each system are highly synchronized, but
the two systems are functionally dissociated.

To quantitatively test for the language-MD functional dissociation, we
compared the average pairwise correlation within the language system (across all
fROI pairs) to the average pairwise correlation within the MD system and the
average pairwise correlation across the two systems (Figure 3) (averages were
computed based on Fisher-transformed correlations; see Materials and Methods).
Specifically, a one-way, repeated-measures ANOVA was carried out to compare
functional correlations among these three levels (“within language”, “within MD”
and “across systems”). Consistent with our qualitative observations, a highly robust
effect was revealed in both conditions (resting-state: $F_{(2,34)} = 62.84$, $p < 10^{-11}$; story-
comprehension: $F_{(2,18)} = 78.56$, $p < 10^{-8}$). In the resting-state condition, post-hoc
pairwise comparisons (Bonferroni corrected, for all that follow) showed that the
average correlation across the two systems ($r = 0.03$, SD 0.11 across subjects) was weaker than the average correlation within the language system ($r = 0.38$, SD 0.13) and within the MD system ($r = 0.41$, SD 0.09) (for both tests, $t_{(17)} > 8.47$, $p < 10^{-6}$).

Similarly, in the story-comprehension condition, the average correlation across the two systems ($r = -0.03$, SD 0.10) was weaker than the average correlation within the MD system ($r = 0.37$, SD 0.11), which was in turn weaker than the average correlation within the language system ($r = 0.49$, SD 0.08) (for both tests, $t_{(9)} > 3.68$, $p < 0.016$).

Within-system and between-system correlations were also compared while controlling for the effect of anatomical distance among fROIs. To this end, we computed the Euclidean distances between each pair of ipsilateral fROIs, based on either their respective center-of-mass coordinates or their point of maximal proximity. Then, for each fROI, its distances from all other ipsilateral fROIs were regressed out from the corresponding functional correlations (separate analyses were performed for the two distance measures). For each fROI, we then tested whether its residual correlations with other fROIs that belonged to its own system were stronger than its residual correlations with fROIs that belonged to the other system. We found that correlations within each system remained stronger than correlations across the two systems. This difference in correlation strength reached significance for all fROIs (Bonferroni-corrected for multiple comparisons) except for two language fROIs in the right hemisphere: the AngG (in both the resting-state and story-comprehension conditions) and the MFG (in the resting-state condition).

**Comparison of the left and right hemispheres**
We next compared, for each of the language and MD systems, the average pairwise correlation within the left hemisphere to the average pairwise correlation within the right hemisphere and the average pairwise correlation across the two hemispheres (Figure 3). Specifically, a one-way, repeated-measures ANOVA was carried out to compare functional correlations among these three levels (“within right hemisphere”, “within left hemisphere” and “across hemispheres”).

For the language system, a significant hemispheric difference was revealed in both the resting-state \( F(2,34) = 28.97, p < 10^{-7} \) and the story-comprehension \( F(1.11,9.99) = 63.95, p < 10^{-5}, \) Greenhouse-Geisser corrected for non-sphericity) conditions. In the resting-state condition, post-hoc pairwise comparisons showed that the average correlation in the left-hemisphere \( r = 0.50, \) SD 0.14 was stronger than the average correlation in the right-hemisphere \( r = 0.40, \) SD 0.13, which was in turn higher than the average correlation across hemispheres \( r = 0.31, \) SD 0.15) (for all tests, \( t(17) > 3.61, p < 0.007 \)). In the story-comprehension condition, post-hoc pairwise comparisons showed that the average correlation within the left-hemisphere \( r = 0.66, \) SD 0.10) was stronger than the average correlation in the right-hemisphere \( r = 0.42, \) SD 0.11) as well as across hemispheres \( r = 0.43, \) SD 0.10) (for both tests, \( t(9) > 7.54, p < 0.001 \)), but the latter two did not significantly differ.

In the MD system, significant hemispheric effects were also revealed in the resting-state \( F(1.22, 20.69) = 8.19, p = 0.007, \) Greenhouse-Geisser corrected) and story-comprehension \( F(2,18) = 13.51, p < 0.001 \) conditions. In the resting-state condition, post-hoc pairwise comparisons showed that the average correlation across hemispheres \( r = 0.37, \) SD 0.10) was weaker than the average correlation within the
left hemisphere ($r = 0.44, SD \ 0.14$) and within the right hemisphere ($r = 0.46, SD \ 0.12$) (for both tests, $t_{(17)} > 4.27, p < 0.002$). Similar results were found for the story-comprehension condition (across hemispheres: $r = 0.33, SD \ 0.13$; left hemisphere: $r = 0.42, SD \ 0.12$; right hemisphere: $r = 0.43, SD \ 0.12$; for both tests, $t_{(9)} > 3.88, p < 0.012$).

**K-means**

*Clustering fROIs*

We clustered fROIs based on the correlations among their respective average BOLD signal time-courses, separately for each subject. For both the resting-state and the story-comprehension conditions, clustering the fROIs into $K = 2$ clusters revealed a clear partition between the language system and the MD system (Figure 4A). Across subjects in the resting-state condition, an average of 14.24(1.44) fROIs out of the 16 candidate language fROIs, or 89%(0.09), were grouped into one cluster; while an average of 16.39(1.32) fROIs out of the 18 candidate MD fROIs, or 91%(0.07), were grouped into a different cluster. Similarly, across subjects in the story-comprehension condition, an average of 14.59(1.00) fROIs out of the 16 candidate language fROIs, or 91%(0.06), were grouped into one cluster; while an average of 16.49(1.80) fROIs out of the 18 candidate MD fROIs, or 92%(0.1), were grouped into a different cluster. A notably inconclusive clustering pattern was only observed for two candidate, right-hemisphere homologue language fROIs, namely the right AngG and MFG. Across subjects in the resting-state condition, these two regions were assigned to the language-dominant cluster only on 57.2%(47) and 51.8%(48) of the clustering solutions, respectively. Across subjects in the story-comprehension
condition, these two regions were assigned to the language-dominant cluster only on 63.5%(28) and 50.5%(34) of the clustering solutions, respectively. Importantly, the separation between language and MD systems did not result from constraining the algorithm to generate exactly two clusters; a similar pattern was obtained for values of $K$ ranging from 3 to 6 clusters, where candidate language fROIs were still clustered with each other more often than with MD fROIs, and vice versa (Figure 5).

The partition of fROIs into a language cluster and a MD cluster was not expected to occur at random, as indicated by a permutation test using surrogate BOLD time-courses (created via phase-shuffling of the original data). Out of the 91 possible pairs of 14 language fROIs (excluding the right AngG and MFG), 86 pairs (94.5%) in the resting-state condition and 91 pairs (100%) in the story-comprehension condition were jointly clustered significantly more often than expected by chance. Similarly, out of the 153 possible pairs of 18 MD fROIs, 137 pairs (89.5%) in the resting-state condition and 153 pairs (100%) in the story-comprehension condition were jointly clustered significantly more often than expected by chance. Conversely, out of 288 possible pairs consisting of a language fROI and a MD fROI, 285 pairs (99%) in the resting-state condition and 288 pairs (100%) in the story-comprehension condition were jointly clustered less often than expected by chance.

**Clustering individual voxels**

When BOLD time-courses from all individual voxels within our fROIs were clustered into $K = 2$ clusters, a "language-dominant" cluster and a "MD-dominant" cluster again emerged. Namely, a high percentage of voxels originating within language
fROIs were all assigned to one cluster (resting-state: 76.59%(9.11); story-comprehension: 83.76%(6.44); averaged across subjects); but this same cluster consisted of much fewer voxels originating within MD fROIs (resting-state: 19.66%(7.67); story-comprehension: 19.38%(9.14)). This cluster was hence language-dominant, with the other cluster showing the opposite, MD-dominant pattern. The difference between the percentage of language voxels assigned to a cluster and the percentage of MD voxels assigned to the same cluster was significant (resting-state: $t_{(17)} = 16.61$, $p < 10^{-11}$; story-comprehension: $t_{(9)} = 15.28$, $p < 10^{-7}$) (Figure 4C).

When we compared, for every given pair of fROIs, the assignments of their constituent voxels into the two clusters, we observed a clustering pattern similar to that reported for fROI-wise clustering. Namely, in both the resting-state and the story-comprehension conditions, a partition between the language and MD systems again emerged (Figure 4B). Thus, the majority of voxels originating in language fROIs were all assigned to the same cluster, whereas the majority of voxels originating in MD fROIs were assigned to the other cluster. The least conclusive clustering pattern was again observed for the right AngG and MFG (candidate, right-homologue language fROIs). Only about half of the voxels originating in these fROIs were assigned to the language-dominant cluster, whereas the other half were assigned to the MD-dominant cluster. Out of the 91 possible pairs of the remaining 14 language fROIs, 86 pairs (94.5%) in the resting-state condition and 91 pairs (100%) in the story-comprehension condition had their voxels jointly clustered significantly more often than expected by chance. Similarly, out of the 153 possible pairs of the 18 MD fROIs, 140 pairs (91.5%) in the resting-state condition and 153
pairs (100%) in the story-comprehension condition had their voxels jointly clustered significantly more often than expected by chance. Conversely, out of the 288 possible pairs consisting of a language fROI and a MD fROI, 286 pairs (99.3%) in the resting-state condition and 287 pairs (99.7%) in the story-comprehension condition had their voxels jointly clustered less often than expected by chance. As was the case for fROI-wise clustering, the general patterns of voxel-wise clustering also did not depend on the choice of $K$ (Figure 5).

**Hierarchical Clustering**

Tree structures (dendrograms) of fROIs constructed based on their correlations revealed a functional architecture dominated by the partition into language and MD systems (Figure 6). In both the resting-state and the story-comprehension conditions, the topmost branching of the tree already separated the set of candidate language fROIs from the set of candidate MD fROIs. Thus, clustering together regions whose distance was shorter than the height of the first branching partitioned the data into language and MD systems. In both experimental conditions, this partition into 2 clusters had the highest modularity value compared to all other possible partitions licensed by the hierarchical tree, suggesting that the functional dissociation between the language and MD systems is the key organizational principle underlying the data.

The organization of fROIs within each of the two systems provides further clues into their functional architecture. This organization could be characterized by using a short distance threshold for clustering regions into small subgroups, and then slowly increasing the threshold so that they merge into larger subgroups. The
functional organization thus revealed was largely similar across the resting-state and story-comprehension conditions.

In the language system, especially during rest, clustering was dominated by hemisphere and, within each hemisphere, regions were clustered according to their lobe. Namely, left temporal regions were clustered together, and only then merged with left frontal regions. Next, these left-hemisphere regions were merged with right-hemisphere regions, which themselves also formed temporal and frontal subsets. The right AngG and MFG were among the last to merge with the rest of the language system (consistent with our K-means results), along with the left AngG and right PostTemp.

In the MD system, clustering was sometimes dominated by hemisphere and sometimes by inter-hemispheric homology. Namely, most frontal and parietal regions tended to merge with each other ipsilaterally before merging across hemispheres. However, the Insula, SMA, ACC and MFG were each first clustered with their contralateral homologue, and only then merged with each other and with the former, "fronto-parietal" subset.

**Discussion**

The findings reported here demonstrate that fMRI BOLD signal fluctuations are strongly correlated among different brain regions of the language system, as well as among different regions of the MD system, but correlations across these two systems are weak or negative. These results are robust, generalizing across two conditions: rest, where signal fluctuations are not driven by an external task, and
story-comprehension, where signal fluctuations are task-evoked; and they arise in
similar form from both hypothesis-driven and data-driven analyses. These data
provide powerful new evidence that each of these systems forms a cohesive,
integrated whole, yet the two systems are functionally dissociated from each other.

Previous studies that used data-driven clustering of voxels across the brain
based on resting-state functional data provided the groundwork for the present
study, and even revealed systems that approximately resemble the language and MD
systems investigated here (Lee et al., 2012; Mantini et al., 2013; Tie et al., 2012; Yeo
et al., 2011). However, the only way to link the clusters that emerged in those
studies to the wealth of knowledge about the functions of different brain regions is
through "reverse inference" based on stereotaxic coordinates (Poldrack, 2006).
Such anatomy-based inferences are particularly challenging for the language and
MD systems, because the mapping between stereotaxic coordinates and functional
regions is degraded by both the high variability across subjects in the anatomical
locations of each region (Amunts et al., 1999; Frost & Goebel, 2012; Juch, Zimine,
Seghier, Lazeyras, & Fasel, 2005; Paus et al., 1996; Tahmasebi et al., 2012;
Tomaiuolo et al., 1999) and the frequent proximity of language and MD regions (e.g.,
Fedorenko, Duncan, et al., 2012). Here we circumvented these problems by
performing our clustering analyses on regions (and voxels) that were defined
functionally within each subject, thereby allowing a direct interpretation of the
resulting clusters in terms of specific functional hypotheses.

The combination of a subject-specific functional localization approach with
an analysis of functional correlations has been previously applied to the ventral
visual pathway (Turk-Browne, Norman-Haignere, & McCarthy, 2010; Zhen, Fang, &
Liu, 2013; Zhu, Zhang, Luo, Dilks, & Liu, 2011) and other regions (Harmelech, Preminger, Wertman, & Malach, 2013; Heinzle, Wenzel, & Haynes, 2012). In fact, this method was used by the first paper to report resting-state functional correlations (Biswal, Zerrin Yetkin, Haughton, & Hyde, 1995). However, no prior study has used this method to study the functional relationship between the language and MD systems. Specifically, while a few prior functional correlation studies did use functional localizers for defining either language (Makuuchi & Friederici, 2013; Newman et al., 2013) or MD (Dosenbach et al., 2007; Seeley et al., 2007) regions, most of these studies used group-level analyses of the localizer data (or coordinates from prior studies; Turken & Dronkers, 2011), again with the potential shortcomings described above. The few studies that did define candidate language regions in individual subjects (Hampson et al., 2002; Morgan, Mishra, Newton, Gore, & Ding, 2009) have focused on small subsets of the language network rather than providing wide coverage of regions of interest.

The current study therefore provides new support for the hypothesis that the language and MD systems are dissociable from each other and are recruited for distinct cognitive processes. Our findings complement prior evidence from neuroimaging studies that used standard functional contrasts (Fedorenko et al., 2011; Fedorenko, Duncan, et al., 2012; Monti & Osherson, 2012; Monti et al., 2009, 2012), as well as evidence from double dissociations in the patient literature (Apperly, Samson, Carroll, Hussain, & Humphreys, 2006; Bek, Blades, Siegal, & Varley, 2010; Broca, 1861/2006; Butterworth, 2000; Happé, Brownell, & Winner, 1999; Klessinger, Szczerbinski, & Varley, 2007; Luria, Tsvetkova, & Futer, 1965;
Remaining questions

A crucial question for further investigation concerns the fine-grained functional organization within the language system and within the MD system. Although each system is highly integrated as indexed by the strong correlations among its constituent regions found here, further functional subdivisions within each system are likely. Indeed, our clustering results already capture some possible subdivisions within each system, and some of these appear to correspond to those suggested in prior studies. Namely, within the MD system, our hierarchical clustering analyses revealed two subsets that may correspond to the previously identified "fronto-parietal" and "cingulo-opercular" networks (Dosenbach et al., 2007; Dosenbach et al., 2006; Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999; Mantini et al., 2013; Nomura et al., 2010; Power et al., 2011). These two networks are hypothesized to be differently recruited for controlling task-relevant cognitive strategies (for a review, see: Power & Petersen, 2013). Within the language system, our hierarchical clustering revealed frontal and temporal subsets in each hemisphere, consistent with previous results from both aphasic patients (e.g., Geschwind, 1970; Gorno-Tempini et al., 2004) and fMRI studies (e.g., Hagoort, 2003; Hagoort, 2005; Snijders et al., 2009; Tie et al., 2012). The current approach of combining functional correlation measures with functional localizers that target specific cognitive functions is likely to prove powerful in further elucidating these hypothesized subdivisions within the language and MD systems.
A second unanswered question concerns the neurobiological significance of functional correlations across brain regions. It has been suggested that these correlations may in part reflect (1) anatomical connectivity (direct and indirect); and/or (2) history of co-activation (Deco & Corbetta, 2011; Deco, Jirsa, & McIntosh, 2010, 2013; for reviews and additional accounts, see: He, Snyder, Zempel, Smyth, & Raichle, 2008; Keller et al., 2011; Matsui et al., 2011; Schölv inck, Maier, Frank, Duyn, & Leopold, 2010; Shmuel & Leopold, 2008). The extent to which the patterns of correlations reported here correspond to direct anatomical connections therefore remains to be discovered. Although evidence from diffusion imaging is generally consistent with resting functional correlation measures (Hermundstad et al., 2013), neither is a perfect measure of structural connectivity (Uğurbil et al., 2013), posing substantial challenges for a definitive answer to this question. Moreover, some functional correlations are critically task-dependent (Hermundstad et al., 2013). Thus, although the functional dissociation between the language and MD systems generalized across resting-state and story-comprehension, it is possible that this dissociation would be modulated under other cognitive states.

Third, although our data indicate that the language and MD regions are independently recruited during cognitive processing, this conclusion need not imply that the two systems can never be engaged simultaneously. Indeed, many previous fMRI studies have reported activations in MD regions during some language processing tasks, especially when such processing is effortful (January et al., 2009; Kuperberg et al., 2003; McMillan, Clark, Gunawardena, Ryant, & Grossman, 2012; McMillan et al., 2013; Meltzer, McArdle, Schafer, & Braun, 2010; Nieuwland, Martin, & Carreiras, 2012; Novais-Santos et al., 2007; Rodd, Davis, & Johnsrude, 2005; Wild
et al., 2012). These findings suggest that the domain-general cognitive-control mechanisms associated with the MD system may play a role in language processing (Fedorenko, under review; Fedorenko & Thompson-Schill, 2014) and, hence, that the MD and language systems may co-activate in some circumstances. Interactions between these two systems, however, may be more pronounced on a fast millisecond-level timescale, and therefore may not be detectable in the BOLD signal fluctuations measured here, given the low temporal resolution of this signal. Thus an important question for future research concerns the frequency, nature, and functional importance of interactions and co-activations of the language and MD systems.

**Conclusions**

Our results support a functional dissociation between the language and MD systems: each system is strongly correlated within itself, but pairs of regions straddling the two systems show weak (or negative) correlations. The robustness of this dissociation across conditions and analyses suggests that it reflects a deep principle of the functional organization of the human brain. Thus, the current data help resolve the controversy in the prior neuroimaging literature (Blumstein & Amso, 2013; Thompson-Schill et al., 2005) in favor of the hypothesis that at least some of the neural mechanisms used for high-level language processing are distinct from those that support other cognitive functions.

**Acknowledgements**
The authors like to thank Susan Whitfield-Gabrieli for valuable advice on connectivity analysis in the CONN toolbox, and Terri Lynn Scott for generating functional regions of interest in native functional space. The authors are also grateful to Anastasia Vishnevetsky, Steve Piantadosi and Ted Gibson for help in constructing and recording the materials for the story-comprehension task, and to Eyal Dechter and Alex Kell for help with the script for the stories experiment. For technical support during scanning, the authors thank Christina Triantafyllou, Steve Shannon, and Sheeba Arnold. We also acknowledge the Athinoula A. Martinos Imaging Center at McGovern Institute for Brain Research, MIT. This research was supported by the following funding sources: grant EY13455 to NK, a grant from the Ellison Medical Foundation to NK, a seed grant from the Simons Center for the Social Brain to NK, and Eunice Kennedy Shriver NICHD Award K99 HD-057522 to EF.

References


Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca’s area. *Current Biology*. 


**Figure legends**

**Figure 1.** Group-constrained, subject-specific fROIs. (A) and (C) show probability maps of the locations of fROIs across subjects, for the language (red) and MD (blue) systems, respectively. Higher color saturation corresponds to a higher number of subjects having a significant activation in the relevant voxel. Apparent overlap between fROIs is only at the group level, not the individual subject level. (B) and (D) show candidate language fROIs (red) and candidate MD fROIs (blue) in the left hemisphere of 3 representative subjects. In all subfigures, dark grey lines demarcate the masks used to constrain the location of fROIs (see section 2.5). All subfigures are in MNI space for illustration purposes only (fROI definition and functional correlation analyses were carried out in the native functional space of each individual subject). Apparent overlap between different fROIs only results from the projection of fROIs onto the cortical surface.

**Figure 2.** Matrices of fROI-to-fROI functional correlations, for (A) the resting-state condition; and (C) the story-comprehension condition. Matrices in (C) and (D) present the same data as (A) and (C) respectively, but show only significant correlations ($\alpha = 0.05$, FDR-corrected). Non-significant correlations are colored in black. The order of fROIs across rows (and columns) follows Table 1, where regions are sorted by system (language, then MD). Within each system, fROIs are sorted by hemisphere (left hemisphere (LH), then right hemisphere (RH)). Thick white lines separate these subsets of fROIs.

**Figure 3.** Comparisons of average correlations within and across systems and hemispheres, for (A) the resting-state condition; and (B) the story-comprehension
condition. Three repeated-measures comparisons are presented. **Left:** comparing the average pairwise correlation within the language system (i.e., across all language fROI pairs), the average correlation within the MD system (i.e., across all MD fROI pairs), and the average correlation between the two systems (i.e., across all pairs of a language fROI and a MD fROI). **Middle:** comparing the average pairwise correlation within the left hemisphere (LH), within the right hemisphere (RH) and between hemispheres, in the language system. **Right:** comparing the average pairwise correlation within the left hemisphere, within the right hemisphere and between hemispheres, in the MD system. Errorbars show standard deviations across subjects. *p < 0.05, **p < 0.01, ***p < 0.001 (Bonferonni-corrected for multiple comparisons).

**Figure 4.** K-means clustering results for the resting-state (left) and story-comprehension (right) conditions, with K = 2 clusters. (A), (D) The average BOLD signal time-course of each fROI was extracted, and the resulting time courses were clustered. In the fROI-to-ROI similarity matrices plotted here, the color of an entry \((i,j)\) for a given pair of fROIs represents the probability (percentage of clustering solutions across subjects and initializations) that the two fROIs would both be assigned to the same cluster. (B), (E) BOLD signal time-courses of all voxels falling within our fROIs were clustered. For each fROI, its "dominant cluster" was then defined as the cluster to which most of the voxels originating within that fROI were assigned. In the fROI-to-ROI similarity matrices plotted here, the color of an entry \((i,j)\) for a given pair of fROIs represents the percentage of voxels in fROI \(j\) that were assigned to the dominant cluster of fROI \(i\) (note that this is not symmetrical).
Percentages are averaged across subjects and initializations. In all matrices (A)-(D) only significant entries are shown (as assessed with a permutation test, based on phase shuffling of the original BOLD time courses; $\alpha = 0.05$, FDR-corrected). Non-significant entries are colored in black. The order of fROIs across rows (and columns) follows Table 1, where regions are sorted by system (language, then MD). Within each system, fROIs are sorted by hemisphere (left hemisphere (LH), then right hemisphere (RH)). Thick white lines separate these subsets of fROIs. (C), (F) Same data as in (B) and (E), respectively. The proportion of "language voxels" and "MD voxels" from each hemisphere that were assigned to each cluster are presented (across the two clusters, bars of the same color add to 100%). Errorbars show standard deviations across subjects.

Figure 5. $K$-means clustering results of functional correlation data as a function of $K$. Conventions are the same as in Figure 4.

Figure 6. Results of hierarchical clustering for (A) the resting-state condition; and (B) the story-comprehension condition. Hierarchical clustering creates a binary tree, with branch length (here, horizontal lines) corresponding to the similarity between fROIs (or sets of fROIs). Above each hierarchical tree, modularity is plotted for all fROI partitions licensed by the tree. Each point on the modularity plot corresponds to a partition generated by drawing an imaginary vertical line from that point through the tree, and clustering together only those fROIs that are merged to the left of this line (fROIs that are merged to the right of the line remain in separate clusters). A sample vertical line is drawn for the maximal modularity,
which corresponds to a partition of the data into two clusters, one consisting of language fROIs and the other consisting of MD fROIs.
(A) Rest: all correlations

(B) Rest: significant correlations

(C) Story comprehension: all correlations

(D) Story comprehension: significant correlations
(A) Rest

(B) Story Comprehension
(A) Rest: fROI-wise K-Means

(B) Rest: voxel-wise K-Means

(C) Rest: voxel-wise K-Means

(D) Story comprehension: fROI-wise K-Means

(E) Story comprehension: voxel-wise K-Means

(F) Story comprehension: voxel-wise K-Means

% voxels assigned

similarity (%)
K = 3 clusters

(A) Rest: fROI-wise K-Means
(B) Rest: voxel-wise K-Means
(C) Story comprehension: fROI-wise K-Means
(D) Story comprehension: voxel-wise K-Means

K = 4 clusters

(A) Rest: fROI-wise K-Means
(B) Rest: voxel-wise K-Means
(C) Story comprehension: fROI-wise K-Means
(D) Story comprehension: voxel-wise K-Means

K = 5 clusters

(A) Rest: fROI-wise K-Means
(B) Rest: voxel-wise K-Means
(C) Story comprehension: fROI-wise K-Means
(D) Story comprehension: voxel-wise K-Means

K = 6 clusters

(A) Rest: fROI-wise K-Means
(B) Rest: voxel-wise K-Means
(C) Story comprehension: fROI-wise K-Means
(D) Story comprehension: voxel-wise K-Means
<table>
<thead>
<tr>
<th>candidate language regions</th>
<th>candidate MD regions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Left Hemisphere</strong></td>
<td><strong>Right Hemisphere</strong></td>
</tr>
<tr>
<td>1 MidPostTemp</td>
<td>9 MidPostTemp</td>
</tr>
<tr>
<td>2 PostTemp</td>
<td>10 PostTemp</td>
</tr>
<tr>
<td>3 MidAntTemp</td>
<td>11 MidAntTemp</td>
</tr>
<tr>
<td>4 AntTemp</td>
<td>12 AntTemp</td>
</tr>
<tr>
<td>5 IFG</td>
<td>13 IFG</td>
</tr>
<tr>
<td>6 IFGOrb</td>
<td>14 IFGOrb</td>
</tr>
<tr>
<td>7 MFG</td>
<td>15 MFG</td>
</tr>
<tr>
<td>8 AngG</td>
<td>16 AngG</td>
</tr>
<tr>
<td></td>
<td>17 IFGop</td>
</tr>
<tr>
<td></td>
<td>18 MFG</td>
</tr>
<tr>
<td></td>
<td>19 MFGOrb</td>
</tr>
<tr>
<td></td>
<td>20 ParInf</td>
</tr>
<tr>
<td></td>
<td>21 ParSup</td>
</tr>
<tr>
<td></td>
<td>22 ACC</td>
</tr>
<tr>
<td></td>
<td>23 Insula</td>
</tr>
<tr>
<td></td>
<td>24 SMA</td>
</tr>
<tr>
<td></td>
<td>25 PrecG</td>
</tr>
<tr>
<td></td>
<td>26 IFGop</td>
</tr>
<tr>
<td></td>
<td>27 MFG</td>
</tr>
<tr>
<td></td>
<td>28 MFGOrb</td>
</tr>
<tr>
<td></td>
<td>29 ParInf</td>
</tr>
<tr>
<td></td>
<td>30 ParSup</td>
</tr>
<tr>
<td></td>
<td>31 ACC</td>
</tr>
<tr>
<td></td>
<td>32 Insula</td>
</tr>
<tr>
<td></td>
<td>33 SMA</td>
</tr>
<tr>
<td></td>
<td>34 PrecG</td>
</tr>
</tbody>
</table>
| Infrequent words          | “Autosomal”  
|                         | “Brunt”       
|                         | “Conjectured” |
| Non-local syntactic dependencies | “The kindly Lord of the Manor who the people had often asked for help...” |
|                          | “The severity of the problem the people faced...” |
|                          | “The water snail that she had discovered a couple of days ago...” |
| Unusual syntactic constructions | “A source of great trouble to the local folk the boar was...” |
|                          | “It was the first huntsman who was...” |
|                          | “Into vapor the water drops that danced in the ocean had been changed...” |
| Temporary ambiguity      | “The huntsman questioned by the Lord...” |
|                          | “The matron understood my idea was something that I was excited about...” |
|                          | “Abby’s mom denied Abby’s version of the story was true...” |