



Published in final edited form as:

Neuroimage. 2015 January 1; 104: 125–137. doi:10.1016/j.neuroimage.2014.10.008.

Language Networks Associated with Computerized Semantic Indices

Serguei V. S. Pakhomov¹, David T. Jones², and David S. Knopman²

¹University of Minnesota Center for Clinical and Cognitive Neuropharmacology

²Department of Neurology, Mayo Clinic, Rochester, MN

Abstract

Tests of generative semantic verbal fluency are widely used to study organization and representation of concepts in the human brain. Previous studies demonstrated that clustering and switching behavior during verbal fluency tasks is supported by multiple brain mechanisms associated with semantic memory and executive control. Previous work relied on manual assessments of semantic relatedness between words and grouping of words into semantic clusters. We investigated a computational linguistic approach to measuring the strength of semantic relatedness between words based on latent semantic analysis of word co-occurrences in a subset of a large online encyclopedia. We computed semantic clustering indices and compared them to brain network connectivity measures obtained with task-free fMRI in a sample consisting of healthy participants and those differentially affected by cognitive impairment. We found that semantic clustering indices were associated with brain network connectivity in distinct areas including fronto-temporal, fronto-parietal and fusiform gyrus regions. This study shows that computerized semantic indices complement traditional assessments of verbal fluency to provide a more complete account of the relationship between brain and verbal behavior involved organization and retrieval of lexical information from memory.

Keywords

generative verbal fluency; semantic memory; latent semantic analysis; semantic clustering; task-free fMRI; semantics

1. Introduction

The question of how conceptual knowledge is represented, organized and accessed from memory continues to be the subject of much research in multiple disciplines including neuroscience (Caramazza & Mahon, 2006; Chan, Butters, Salmon, & McGuire, 1993;

© 2014 Elsevier Inc. All rights reserved.

Corresponding author: Serguei V.S. Pakhomov, PhD, Associate Professor, University of Minnesota, 308 Harvard St. SE, Minneapolis, MN 55455, Tel. +1 612-624-1198, Fax. +1 612-625-9931, pakh0002@umn.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Mahon & Caramazza, 2009; Patterson, Nestor, & Rogers, 2007; Pytkkanen, Llinas, & Murphy, 2006), neuropsychology (Salmon, Butters, & Chan, 1999; Troster et al., 1998; Weber, Thompson-Schill, Osherson, Haxby, & Parsons, 2009), psycholinguistics (Levitt, 1989; Tverski & Hemenway, 1984), and computer science (G. Miller & Fellbaum, 1991; Rada, Mili, Bicknell, & Blettner, 1989; P. Resnik, 1999). Following Tulving (1972) this field of inquiry has drawn a distinction between a cognitive system that represents temporally dated events (episodic memory) and a system that constitutes a mental thesaurus of symbols that indexes information about cognitive referents and relations holding between them (semantic memory). The two types of memory have since been demonstrated to have different, albeit interdependent and interconnected, underlying neural mechanisms (Cabeza & Nyberg, 2000; Prince, Tsukiura, & Cabeza, 2007; Tulving, Kapur, Craik, Moscovitch, & Houle, 1994).

The test of generative semantic verbal fluency (SVF) is an instrument widely used to elicit responses from subjects when studying semantic memory, executive function, and language. On this test, one is asked to say as many words as possible in one minute that denote objects belonging to a certain semantic category (e.g., animals, fruits, vegetables, tools). The performance on the SVF test is typically measured by counting the number of correct responses (SVF score). Lower SVF scores have been widely reported in patients with various stages of Alzheimer's disease (AD) dementia and Mild Cognitive Impairment (MCI) (Chan, Salmon, & De La Pena, 2001; Ober, Dronkers, Koss, Delis, & Friedland, 1986; Rosen, 1980; Troyer, Moscovitch, Winocur, Leach, & Freedman, 1998). Furthermore, performance on SVF tests often shows early and more pronounced decline relative to other language, attention, and executive abilities (see Lezak, (2004) and Henry *et al.*, (2004) for review). The traditional SVF scores reflecting the number of words generated on this task assess performance consisting of several abilities including but not limited to semantic memory. Structural and functional neural imaging studies found that lower SVF scores are associated with lesions and atrophy in the anterior and inferior left temporal lobe regions as well as impairment in fronto-temporal connectivity (Libon et al., 2009). Evidence collected in fMRI studies supports the involvement of multiple frontal, temporal and posterior areas during word generation tasks such as SVF that include (but are not limited to) the left inferior frontal gyrus, the left inferior temporal gyrus, the hippocampus, the left superior occipital gyrus, and the left inferior medial parietal lobe (Birn et al., 2010; Wheatley, Weisberg, Beauchamp, & Martin, 2005). Furthermore, the basal ganglia exerts inhibitory control over motor, cognitive/executive, and affective systems which would make it an important part in a variety of tasks including aspects of the SVF performance. The involvement of multiple brain areas in generating words on SVF tests suggests that verbal behavior resulting from this task is supported by multiple neural mechanisms. However, the correspondence between various aspects of the generative verbal fluency behavior and distinct neural mechanisms is currently much less clear because of the difficulty involved in isolating quantifiable characteristics of the responses produced during SVF testing.

The overarching objective of the current study was to investigate the utility of a computational linguistic approach to capturing distinct components of verbal behavior manifest on SVF tests and to relate them to the underlying brain networks identified with

task-free fMRI. Using task-free fMRI in connection with behavioral performance on a specific task is well-motivated by the fact that the patterns of temporally correlated low-frequency fluctuations observed in a resting state constitute the brain's intrinsic architecture that predicts the brain's functional responses to stimulation (Keller et al., 2011).

1.1 Clustering of Responses on SVF Tests

Optimal performance on SVF tests depends to a large extent on how well semantic information is organized into conceptually related clusters and whether the person is able to use an efficient strategy that accesses these clusters during the test (Estes, 1974; Hodges & Patterson, 1995; Laine, 1988). The size of semantic clusters and the efficiency of switching between clusters have been found to have different neuroanatomical correlates. Semantic cluster size was found to be associated with the left temporal lobe function, whereas the processing associated with switching was associated with the function of the frontal lobe (Rich, Troyer, Bylsma, & Brandt, 1999; Troyer, Moscovitch, & Winocur, 1997; Troyer, 2000).

While manual assessment of clustering and switching behavior has proven to be useful, it has traditionally relied on subjective evaluations of semantic similarity between at least two (Rich et al., 1999; Troyer et al., 1997, 1998) or three (Laine, 1988) adjacent words to define semantic clusters. For example, the qualitative assessment proposed by Troyer *et al.* (1997) relies on manual categorization of words produced on the SVF test (e.g., Zoological Categories, Human Use, and Living Environment) with further more fine-grained subcategorizations (e.g. living environment category composed of African, Australian, Arctic/Far North, Farm, North American and Water Animals). In addition to their subjectivity, these manual qualitative approaches are time consuming and are difficult to implement and standardize, which may be responsible for some of the conflicting results obtained with these methods in studies of Alzheimer's disease noted in previous work (Raoux et al., 2008).

Independently of these efforts, a number of fully automated approaches to representing the degree to which any two words in a given language are semantically related have been developed in the field of computational linguistics based on lexical databases such as WordNet, as well as corpora of text (Pedersen, Pakhomov, Patwardhan, & Chute, 2007; Rada et al., 1989; P. Resnik, 1999). Many of these approaches utilize variations on a technique called Latent Semantic Analysis (LSA: (Landauer & Dumais, 1997)), a variant of factor analysis designed for representing lexical semantics. In addition to the LSA approach to semantic representation, several other alternatives have been proposed to model how semantic information is represented in the brain including neural networks (McClelland & Rogers, 2003), Random Indexing (Kanerva, 2009), Latent Dirichlet Allocation (LDA) modeling (Blei, Griffiths, Jordan, & Lafferty, 2003), and distributional memory models (Baroni & Lenci, 2010; Baroni, Murphy, Barbu, & Poesio, 2010).

The application of LSA to semantic representation is described in detail in the Methods section. In brief, LSA relies on the co-occurrence of words in a large corpus of text consisting of various types of discourse including newspaper articles, books, speeches and other sources of typical word usage to represent the semantic content of a word or a term as

a set of co-occurrence counts with other words used in the same context. These semantic representations can then be directly and automatically compared to each other to assign a numeric value indicative of the strength of semantic relatedness between them. Apart from improved scalability and objectivity as a result of automation, these computational approaches allow quantification of semantic relations on a continuous rather than a categorical scale which allows us to: a) directly control and systematically vary how measures such as the cluster size, for example, are calculated, and b) develop new semantic indices not possible with categorical judgments. We have previously reported on applications of these computerized semantic indices, either calculated from WordNet, a large lexical database (Pakhomov, Hemmy, & Lim, 2012), or from a corpus of text (Pakhomov & Hemmy, 2013). In other prior work, computational models of word meanings derived from a very large corpus of text have been demonstrated to predict neural activation patterns observed with fMRI (Mitchell et al., 2008). These findings were based on representations for concrete nouns and thus provide a strong motivation for using distributional semantic approaches to represent the meaning of words produced in response to a verbal fluency task.

The mechanisms underlying *semantic* memory are negatively affected by aging (Meinzer et al., 2009) and are the target of several types of neurodegenerative diseases including the semantic variant of fronto-temporal dementia (Grossman, 2002; Hodges et al., 2004; Knopman et al., 2008) and the Alzheimer's disease (AD) dementia (Hodges & Patterson, 1995). In our previous work, we found that computerized semantic indices were sensitive to clinical differences between mild cognitive impairment (MCI) and AD dementia (Pakhomov et al., 2012), and could be used to estimate future risk of developing dementia in healthy individuals (Pakhomov & Hemmy, 2013). The current study relies on a sample consisting of cognitively normal individuals as well as MCI and AD dementia patients in order to investigate the relationship between SVF performance and functional connectivity in the language network 'at rest' assessed with task-free fMRI. Thus, the disease status is used in this study as a naturally occurring research instrument that modulates both the behavioral performance on SVF tests and the connectivity of brain networks that underlie verbal behavior. Therefore, we expect that the degree of differential impairment in the connectivity of functional brain networks that is characteristic of AD dementia (Stam et al., 2009; Stam, Jones, Nolte, Breakspear, & Scheltens, 2007; Supekar, Menon, Rubin, Musen, & Greicius, 2008) will introduce detectable systematic variability into both behavioral and neural measurements. The specific aims of this study were: a) to confirm previously found neural correlates of the performance on the SVF test, and b) to determine if the new automated semantic clustering indices derived with the LSA-based computational linguistic approach are associated with connectivity in areas distinct from those related to the traditional SVF scores. Initially, our prediction was that the computerized semantic indices were associated with roughly the same brain networks as those associated with traditional SVF scores; however, while we did find some overlap between measures, we also found that they were associated with clearly distinct networks.

2. Methods

2.1 Participants

A random target sample of 60 participants (mean age = 72.46; SD = 10.8) was obtained from the Mayo Clinic Alzheimer's Disease Research Center (MCADRC). Of these participants, at the time of neuropsychological testing, 21 had a clinical diagnosis of probable AD dementia (DSM-IV/NINCDS-ADRDA criteria; (American Psychiatric Association, 1994; McKhann et al., 1984)), 20 had a clinical diagnosis of MCI (Petersen, 2004) with or without an amnesic component, and 19 were cognitively normal elders (CN). The final study sample consisted of 49 participants (11 CN, 17 MCI and 21 AD dementia) that were selected from the target sample of 60 participants based on availability of good quality task-free functional MRI scans (TF-fMRI). The demographic characteristics of this sample are reported in Table 1.

All participants, or appropriate surrogates, provided written informed consent for participation. The Mayo Clinic Institutional Review Board approved the study and the consenting processes

2.2 Cognitive Assessments

All participants underwent a neuropsychological test battery that included the SVF test ("animals" category) on which they were asked to name as many animals as they could think of in 60 seconds. The responses were recorded on paper by a trained psychometrist and subsequently converted to electronic form for computerized analysis. The traditional SVF test score was calculated as the number of correct words excluding repetitions, intrusions and perseverations. The participants were also administered the Short Test of Mental Status (STMS: (Kokmen, Smith, Petersen, Tangalos, & Ivnik, 1991)). The STMS battery items test several cognitive domains: orientation, attention, immediate recall, arithmetic, abstraction, construction, information and recall after a 3-minute delay. This test is designed specifically for rapid (5 min) assessment of cognitive status at bedside and has been demonstrated to be more sensitive than the Mini-Mental State Examination (MMSE) in detecting deficits in individuals with normal cognition at baseline that later developed MCI or AD dementia (Tang-Wai et al., 2003). In addition to the STMS, Clinical Dementia Rating (CDR) scores were also obtained for all participants. The diagnoses of MCI and AD dementia were made during consensus conferences including neurologists, neuropsychologists and nurses and took into account the full neuropsychological test battery, the neurological assessment and the views of the family informants, as obtained by the examining neurologists and nurses.

2.4 Automated Semantic Relatedness Computation

To compute semantic relatedness between pairs of words, we relied on the LSA approach to semantic representation (Landauer & Dumais, 1997; Landauer, 2006). In order to compare the meanings of any given pair of words, we first represent the semantic content of each word as a set of other words found in the same context as the target word in a collection of texts. Since in this study, we are focused on the animal category SVF test, we experimented with the Wikipedia entries for animals, readily available electronically, as a source of textual co-occurrence information. For example, the meaning of the word "tiger" was represented as

a set of all other words that were found in the Wikipedia entry for “tiger” after exclusion of function words (e.g., “the”, “he”, “she”, “it”, “on”, “is”, etc.) including: “panthera”, “tigris”, “largest”, “cat”, “species”, “most”, “recognizable”, “feature”, “pattern”, “dark”, “vertical”, “stripes”, “reddish”, “orange”, “fur”, “Russia”, “Bangladesh”, “India”, “siberian”, “asia”, among others. Similarly, the meaning for the word “lion” contains words “panthera”, “leo”, “four”, “big”, “cats”, “genus”, “panthera”, “second”, “largest”, “living”, “cat”, “tiger”, “currently”, “exist”, “subsaharan”, “Africa”, “asia.” Clearly, the Wikipedia entries also contain many words irrelevant to comparing the meanings of animal names and constitute “noise” that needs to be removed automatically prior to comparing semantic representations distributed across Wikipedia entries. Furthermore, in some cases, it is useful to take advantage of the fact that a given pair of words may not appear in the same context but may still be linked through their co-occurrence with a third word. For example, the words, “tigris” and “leo” may never occur in the same Wikipedia entry but the word “tigris” co-occurs with “panthera” and so does the word “leo”, thus forming a latent semantic association. We used LSA to take advantage of these latent associations between words never occurring in the same context.

LSA is a computational technique that operates by constructing a co-occurrence matrix for all words found in a given corpus of text (e.g., Wikipedia) and applying a variant of principal components analysis to filter out irrelevant words (dimensions) through singular value decomposition and to identify latent semantic associations between the words (Landauer & Dumais, 1997; Landauer, 2006). For the current study, we used a subset of the Wikipedia corresponding to 254 articles about the animals found in the SVF test responses as the source of co-occurrence information to construct semantic representations of animal names. The motivation for using a subset of Wikipedia stems from the fact that corpus-based statistical approaches to representing lexical semantics are subject to potential noise in the data and may result in some cases in “discovering” spurious relations. We took a number of steps previously reported elsewhere (Pakhomov & Hemmy, 2013; Pakhomov et al., 2012) to automatically reduce the amount of noise and to optimize the parameters used in LSA computation and subsequent clustering of responses. In order to eliminate potentially spurious associations, we limited Wikipedia to a subset consisting only of animal names that occurred in the responses. Thus, the original co-occurrence matrix was composed of 254 animal names (Wikipedia documents) and 17,763 context words. While this is a departure from the standard LSA approach in which all available text from a corpus would be used, prior work by Pereira and colleagues (Pereira, Botvinick, & Detre, 2013) on the same dataset that was developed by Mitchel et al. (2008) demonstrated that reasonable topic-model representations of semantic space of concrete concepts can be obtained from a relatively small (~ few thousand entries) subset of Wikipedia definitions.

Apart from the choice of text to be used for LSA, another source of variability has to do with selecting the most optimal dimensionality of LSA vectors. To select the number of dimensions we relied on calculating the proportion (share) of the sum of singular values for the first N dimensions in the LSA matrix to the total sum of singular values for all dimensions. The number of dimensions for LSA computation is difficult to estimate a priori. As described in detail in (Quesada, 2011, p. 82), unlike other related techniques such as multi-dimensional scaling, LSA does not currently have an internal criterion or a theoretical

way of determining the most optimal number of dimensions. Therefore, we determined the optimal dimensionality by using an external criterion – a small independent dataset consisting of pairs of animals manually assessed for semantic similarity as follows. As described in more detail in our previously reported work (Pakhomov & Hemmy, 2013), to control for the effects of source text selection and to find the optimal number of dimensions, we used a dataset developed by Weber and colleagues (Weber et al., 2009) that consists of nine animal names and all possible pairs ($n=36$) of these nine animals that were presented to participants in a series of behavioral and imaging experiments. The participants were asked to judge photographic images of these pairs of animals for similarity. While holding the source data constant (the Wikipedia corpus), we varied the proportion of the sum of singular values for the first N dimensions to the total sum of singular values for all dimensions. We experimented with several values for this criterion in the range between 0.05 and 0.50 and found that the optimal threshold was around 0.20. This threshold resulted in reducing the original word by document matrix to 14 dimensions, yielding a peak correlation of 0.65 between the relatedness scores obtained with this LSA method and the manually derived Weber's reference standard. The previously mentioned work by Pereira et al. also suggests that a relatively small number of dimensions ($\sim 25 - 100$) is needed to model representations of concrete objects using Latent Dirichlet Allocation (LDA) which, similarly to LSA, models the distributions of words over documents in order to “learn” semantic features that best represent them (Pereira, Botvinick, & Detre, 2013).

Applying LSA to Wikipedia entries for all animal names that were produced by the participants in our study resulted in representing each animal name as a set of a number of latent semantic dimensions. We subsequently used the resulting matrix to compare pairs of animal names represented as vectors in this multi-dimensional space by computing the cosine of the angle between the semantic vectors that represented the meaning of each animal name. The cosine values range from -1 (180 degree angle between vectors, i.e., they are pointing in opposite directions), to zero (90 degree angle - the two vectors are orthogonal), to 1 (zero degree angle - the vectors are pointing in the same direction). Thus, these cosine values can be interpreted in terms of semantic relatedness between the words that the vectors represent. We used the cosine values to compute measures of clustering and cumulative semantic relatedness as described in the next two sections – 2.5 and 2.6.

2.5 Automatic Determination of Clusters and their Size in SVF Responses

Based on the semantic relatedness tools described in the previous section, we developed an automated approach that follows the clustering and switching analysis introduced by Troyer et al. (1997) as closely as possible. The one significant departure from Troyer's approach in the current implementation is that the qualitative human judgments of whether any given pair of words belongs to the same semantic category are replaced by quantitative semantic relatedness scores with the cutoff threshold for the cosine values set to 0.90 to identify very closely related pairs (i.e., their semantic vectors point almost in the same direction). Thus, if the relatedness value between two words exceeds this threshold, the words (or to be more precise, the animal senses of these words) are treated as belonging to the same cluster. The rest of clustering computation was the same as described by Troyer et al. (1997), repeated here briefly for convenience. In order to be counted as a cluster, all words in a given

sequence of words found in a test response had to be closely related (i.e., have pairwise relatedness values above the predetermined threshold – 0.90). For example, a sequence of w_1 , w_2 and w_3 would be considered a cluster if w_1 is closely related to w_2 , w_2 is closely related to w_3 , and w_3 is closely related to w_1 . The size of clusters was discounted by one – a single word cluster would have the size of zero, a cluster of two words would have the size of one, and so on; however, the total number of clusters used in the denominator for calculating the mean cluster size (MCS) included single word clusters. It is important to note that we followed Troyer’s methodology to make our results interpretable in light of the prior work; however, automated computation of semantic relatedness enables further experimentation with modified approaches that we reserved for future work¹. For clarity of presentation we will refer to LSA-based automated computation of mean cluster size as MCS and manually computed cluster size based on Troyer’s guidelines as mMCS.

The network graph in Figure 1 illustrates the connectivity between names for the top 50 most frequent animals found in the study participants’ responses. This network was constructed by first computing a fully connected adjacency matrix representing pairwise semantic relatedness values between all 50 animal names. We then applied the same 0.90 threshold as was used in the determination of semantic clusters in participants’ responses to retain only the strongest associations. The resulting network was visually represented with an edge-weighted layout (Cytoscape software package, version 2.8.1, <http://www.cytoscape.org/>) in which the length of the edges is proportionate to cosine values between retained word pairs.

2.6 Measuring Cumulative Semantic Relatedness

In the present study, in addition to clustering and switching, we assessed the *cumulative* approach to computing relatedness over the entire SVF sample. This approach consists of making automatic semantic relatedness measurements between all pairs of words produced during the test of verbal fluency regardless of the order in which they appear in the test sample. The pair-wise measurements are then averaged over the entire sample to obtain a cumulative relatedness (CuRel) score that represents a measure of semantic diversity of the response. We used the word “cumulative” to describe this approach to reflect the fact that the computation of the mean “accumulates” values across all pairs of words even though this approach is not purely additive. This approach is illustrated in Figure 1 in which each test sample is represented as an $N \times N$ matrix (N = number of words generated on the test) with the cells containing pairwise relatedness values with the exception of the on-diagonal cells. Since the same words are used both in rows and columns, these matrices are symmetrical and thus we do not need to compute values for half of the pairs (represented with blank cells). The example in Figure 2 illustrates the computation of CuRel in two samples. Sample 1 is from a 77-year-old male diagnosed with multiple-domain MCI with an amnesic component (STMS: 36, CDR: 0.5). Sample 2 is from an 80-year-old male diagnosed with probable AD dementia (STMS: 30, CDR: 0.5). Theoretically, CuRel score may represent the degree of semantic diversity in the meaning of words produced on the SVF test. Higher

¹The methods for computing MCS and CuRel described in this manuscript as well as additional approaches to measuring clustering and switching in semantic and phonemic fluency test responses have been made publicly available as a Python package (<https://ypi.python.org/pypi/VFClust/0.1.1>).

CuRel scores may index a tendency to produce exemplars that are more closely related and are thus less semantically diverse. Lower CuRel score may indicate greater diversity in exemplars and a tendency to search across subcategories.

2.7 Task Free Functional MRI

Given that performance on the SVF test is multidimensional with various measures of semantic relatedness each likely capturing a distinct dimension of overall performance, it would be expected that the neural substrates supporting these measures would be distinct, at least to some extent, for each measure as well. However, it is also possible that the various aspects of overall performance on this test are controlled by the same or overlapping mechanisms. Therefore, we investigated the association between semantic relatedness measures and the brain systems-level organization of the language network using task-free fMRI.

Task-free fMRI data were acquired using a General Electric 3T Signa HDx scanner, 8 channel head coil, gradient EPI, TR = 3000 ms, TE = 30 ms, 90° flip angle, 21 cm field of view, 64 × 64 in-plane matrix, slice thickness 3.3 mm without gap, and 103 or 113 volumes were obtained. Subjects were instructed to keep their eyes open during scanning. All task-free fMRI data sets with greater than 3 mm of translational movement, 3° of rotational movement, or that failed visual inspection for obvious artifacts were excluded from analysis. Preprocessing and data analysis were performed utilizing a combination of the Statistical Parametric Mapping (SPM5) software (<http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>) (Wellcome Department of Cognitive Neurology, University College London, UK), the Resting-State fMRI Data Analysis Toolkit (REST) v1.5 (<http://www.restfmri.net>) (Song et al., 2011), Data Processing Assistant for Resting-State fMRI (DPARSF) v2.0 (<http://www.restfmri.net>) (Chao-Gan & Yu-Feng, 2010), group ICA of fMRI toolbox (GIFT) software v2.0c (<http://icatb.sourceforge.net>) (Calhoun, Adali, Pearlson, & Pekar, 2001), and in-house developed software implemented in MATLAB v7.11 (Mathworks Inc., Natick, MA).

Preprocessing steps included discarding the first 3 volumes to obtain steady state magnetization (sequences with 113 volumes were also truncated so all sequences had 100 remaining volumes for analysis), realignment, slice time correction, normalization to SPM5 EPI template, smoothing with 4 mm full-width half maximum Gaussian kernel, linear detrending to correct for signal drift, and 0.01–0.08 HZ bandpass filtering to reduce non-neuronal contributions to BOLD fluctuations. In addition, regression correction for spurious variables included rigid body transformation motion effects, global mean signal, white matter signal and cerebral spinal fluid signal (Fox, Zhang, Snyder, & Raichle, 2009; Weissenbacher et al., 2009).

The language network was back reconstructed for each subject using the spatial-temporal dual regression (STR) method, as implemented in the GIFT software package (Calhoun et al., 2001), with scaling of the parameter estimates of functional connectivity to z-scores. The resulting spatial maps contain voxel-wise information about the spatial location and magnitude of language network connectivity at the individual subject level. Independent component analysis (ICA) applied to task-free fMRI identifies functionally connected brain

networks in a model-free manner by estimating spatially independent patterns from linearly mixed time courses of fMRI data. Multi-subject or group ICA (GICA) methods have been developed in order to draw inference about group data using ICA (Erhardt et al., 2011). Using the GICA approach, an aggregate estimate of intrinsic connectivity networks can be obtained. However, subject-level maps need to be back-reconstructed using the group-level solution to initialize the STR method to estimate a subject-level solution (Beckmann, Mackay, Filippini, & Smith, 2009). We have recently developed a well characterized high-dimensional GICA atlas (Mayo Clinic Study of Aging Functional Connectivity Atlas) using a large ($n = 892$) population-based sampling of cognitively normal elderly individuals (Jones et al., 2012). This GICA atlas can be used to initialize the STR method in a manner unbiased by the connectivity profiles of the sample of subjects under investigation.

The STR method is a two-step process to derive subject-level parameter estimates of the intrinsic connectivity network (ICN). The first step uses spatial-maps from group average estimates of ICNs as spatial regressors against a single subject's 4-dimensional dataset that produces a time-series for each subject. This time-series is then used in the second step as a temporal regressor against the same subject's 4-dimensional dataset producing subject-level spatial maps of ICNs. The aggregate GICA language network template from the high-dimensional ICA Mayo Clinic Study of Aging (MCSA) Functional Connectivity Atlas was used for the back-reconstruction (available at http://mayoresearch.mayo.edu/mayo/research/jack_lab/supplement.cfm) (Jones et al., 2012). Each of the subjects' back-reconstructed language networks were entered into a multiple regression SPM5 analysis with age and gender as nuisance covariates and SVF, CuRel, or MCS as the covariate of interest. A one-sample t-test of the back-reconstructed language networks was performed to visualize the spatial extent of the entire language network at the group level (Figure 3). For each of the covariates of interest, a linear regression analysis was performed and considered significant at a FWE corrected cluster level significance $p < 0.05$ (voxel-wise $p < 0.05$, $k > 93$). Each of the identified significant clusters was further characterized using the Anatomy toolbox (Eickhoff *et al.*, 2005) and the MCSA high dimensional ICA ROI atlas (Jones et al., 2012). The average functional connectivity in the most significant clusters for each contrast was extracted from every subject and correlated with the respective covariate of interest in order to better characterize the direction of the relationship between language network connectivity and behavioral performance on a language task.

2.8 Statistical Analysis

Correlations between continuous variables were estimated with Pearson correlation. Differences between group means were tested for statistical significance using one-way ANOVA with post-hoc tests carried out using the Tukey Honestly Significant Differences (HSD) method. Linear regression analysis was performed by including age, sex, and years of education as covariates along with the predictive covariate of interest. Comparisons between manual and automatic determination of clusters in SVF responses were performed based on manual clustering guidelines originally proposed by Troyer and colleagues (Troyer et al., 1997). All statistical tests assumed the significance threshold of 0.05. All statistical computations were carried out using the R statistical package (version 2.14.1) and software designed specifically for brain imaging analysis described in the previous section.

3. Results

3.1 Comparison of Automated and Manual Clustering

Comparing MCS calculated with the LSA approach to mMCS - cluster size determined manually by following Troyer's guidelines – resulted in a moderate correlation of 0.57 ($n = 49$, $p < 0.001$).

3.2 Differences Between Diagnostic Groups

Demographic and clinical characteristics of the participants in this study are presented in Table 1. Significant differences in SVF score means between AD dementia and CN (adjusted p -value < 0.001), and between MCI and CN groups (adjusted p -value = 0.016) were observed with higher SVF scores in CN than in AD dementia and MCI groups. MCI participants were less impaired (higher SVF score) than AD; however, while this difference was in the expected direction, it only approached significance (adjusted p -value = 0.054). On STMS and CDR, post-hoc analyses revealed that both CN and MCI participants were significantly less impaired than AD (adjusted p -value < 0.001); however, the differences between CN and MCI on the STMS and CDR scales were not significant. The CuRel scores were highest in the MCI group followed by a tie between AD dementia and CN; however, these differences were not significant either. A similar pattern of results was observed for STMS, showing that CN and MCI participants performed significantly better than AD, with MCI scores being between CN and AD but the differences failing to reach statistical significance.

The traditional SVF score was not correlated with either MCS ($r = 0.27$, $p = 0.06$) or CuRel ($r = -0.09$, $p = 0.52$) measures. MCS and CuRel were also uncorrelated ($r = 0.16$, $p = 0.28$). We estimated three regression models with the STMS score as the dependent variable and SVF, MCS, or CuRel score as the independent variable, respectively. Each model also included sex, age and education level of the participants as covariates in addition to the SVF, MCS, or CuRel score. The model with the SVF score showed that it was a significant predictor of STMS ($p < 0.001$); however, the remaining models did not show significant associations between the LSA-based scores (MCS or CuRel) and STMS.

3.3 Associations Between Language Brain Network Connectivity and Semantic Clustering

The back-reconstructed language network recapitulated the typical regions of positive and negative functional connectivity found in the ICN shown in Figure 3. The spatial topography of ICNs represents the average composition of dynamic binary brain configurations (Jones et al., 2012), consisting of regions of positive synchrony in brain areas related to a specific cognitive task and regions of negative synchrony (i.e., anticorrelations) in brain areas which may serve related, yet competing, functions (Fox et al., 2009). Hence, the optimal brain organization that supports language-related tasks, captured by analysis of ICNs in both task-related and task-free experimental paradigms, relates to the system's ability to maintain appropriate positive and negative functional connectivity of these regions, shown in Figure 3.

In this study, we examined the associations between one of each of the three covariates of interest (SVF, MCS, and CuRel) and both positive and negative regions of functional connectivity in the back-reconstructed language network (Table 2 and Figure 3). We looked for associations between performance on semantic clustering measures with functional connectivity of individual language network regions giving rise to four possibilities: loss of positive functional connectivity, loss of negative functional connectivity, gain of positive functional connectivity, or gain of negative functional connectivity. In general, poorer performance on semantic indices was associated with a loss of either positive or negative functional connectivity highlighting the importance of both strong positive and strong negative connectivity for high performance. The only exception being the gain of positive functional connectivity seen in the medial frontal lobe associated with lower SVF scores discussed below.

The following results are organized in terms of the comparisons between the back-reconstructed language network (see Figure 3 showing the sign of functional connectivity of various regions with the rest of the language network) and the relationship of specific language network areas to the covariates of interest (see Figure 3 showing the direction of the correlations between functional connectivity in language network areas and the covariates). It is important to note that yellow and blue colors in Figure 3 represent a different phenomenon from the colors in Figure 4. The colors in Figure 3 index areas of positive and negative functional brain network connectivity, whereas colors in Figure 4 positive and negative correlations with semantic indices.

Better performance measured by the traditional SVF score was most strongly correlated with the left middle temporal gyrus and the basal ganglia (Table 2). Both of these areas display positive functional connectivity with the rest of the language network (see Figure 3, yellow and orange areas). The negative functional connectivity of the medial prefrontal cortex (see Figure 3, blue), which is part of the anterior default mode network (aDMN), was correlated with poorer performance measured with the SVF score (shown in Figure 4a, blue).

The MCS measure was found to be differentially associated with the middle temporal gyrus, superior parietal lobule, and fusiform gyrus. The middle temporal region had positive functional connectivity with the rest of the language network (Figure 3, orange and yellow). Greater positive connectivity in this region was associated with larger semantic clusters (Figure 4b, yellow); however, larger semantic clusters were also associated with greater negative connectivity of the left superior parietal lobule and fusiform gyrus (Figure 4b, blue and Figure 3, blue).

Lower CuRel scores were associated with loss of functional connectivity in the left angular gyrus and left frontal operculum areas (Figure 4c, blue). Both of these areas are shown in Figure 3 (yellow and orange) to be positively connected with the rest of the language network, thus the detrimental association with CuRel scores in this case is loss of positive functional connectivity. Higher CuRel scores were associated with greater negative functional connectivity of the ventral DMN areas (right precuneus and retrosplenial posterior cingulate cortex, Figure 4c, yellow). At first glance it may seem counterintuitive that larger CuRel scores are associated with detrimental loss of positive and negative

functional connectivity, particularly in light of the fact that CuRel means do not differ significantly across diagnostic groups. In order to further elucidate the relationship between CuRel scores and behavioral performance, we examined the association between CuRel and SVF scores by subgroups with higher and lower mental status using the STMS score threshold of 34 to separate participants into these groups. The motivation for using STMS score cutoff of 34 stemmed from findings of a previous large study that examined the utility of using STMS to predict MCI (Tang-Wai et al., 2003) that reported a mean of 34.2 for patients with MCI diagnosis. The results of this examination are shown in Figure 5 and are juxtaposed with the associations in Figure 4c, repeated in Figure 5 for convenience. These results show that there is a significant negative association between CuRel and SVF performance ($r = -0.52$, $p = 0.022$) in the higher functioning group (Figure 5a, left). No association with SVF performance ($r = 0.01$, $p = 0.98$) was found for the lower-functioning group. The direction of the associations between CuRel and loss of negative and positive functional connectivity in the left angular gyrus and the right precuneus, respectively, was preserved.

4. Discussion

4.1 General Discussion

In this study, we demonstrated that distributed semantic representations of animal names constructed using LSA from a small subset of a large online encyclopedia can be used to characterize verbal behavior elicited with SVF tests and investigate distinct brain mechanism(s) that support generative semantic fluency. While two-thirds of our study sample came from a clinical population, our focus was not clinical. In this study, the disease process that resulted in MCI and AD dementia in the participants was used as a naturally occurring research instrument that modulated the connectivity of brain networks involved in verbal responses to the generative semantic verbal fluency task. The results of this study confirm the original proposal by Troyer et al. (1997) that cognitive performance on the SVF test is multidimensional and that the traditional SVF scores do not account for every aspect of performance on this test. We have developed a computational linguistic approach that can help assess multiple behavioral aspects of SVF testing in addition to the SVF score in a standardized and reproducible fashion, and improve the test's explanatory power with respect to the underlying brain mechanisms involved in various aspects of this task. Our approach leverages the information contained in a new generation of social media resources (such as the Wikipedia), that are constantly being created and updated through Internet-mediated interactions among large numbers of people. Thus these resources can be used as a valuable source of information to represent distributional semantic properties of words.

Troyer et al.'s (1997) procedure is widely used in research; however, it has certain limitations. It relies on an experimentally defined shallow hierarchy (at most 3 levels) with three top-level categories (Living Environment, Human Use and Zoological Categories) and several sub-categories with multiple inheritance (the same animal name may appear in more than one category). The subcategories in this classification scheme were defined based on patterns experimentally observed (unfortunately, the authors did not specify what empirical patterns were used to determine category membership) in the tests administered to 95

healthy Canadian adults. Thus manual clustering using this categorization scheme assumes that the subject conceptualizes the animal kingdom in the same way as this relatively small and specific study population. For example, Troyer's categorization places the word "seal" under "Arctic/Far North Animals", which makes sense from the standpoint of someone living in Canada; however, seals are also found in many places other than the Arctic. This is a trivial example used here only to illustrate the arbitrary nature of picking the dimension (among the three available dimensions exemplified by the top level categories) along which an animal can be categorized.

In contrast to Troyer's original manual approach, the LSA-based automated clustering approach that we propose results in semantic relatedness assessments between words along multiple dimensions derived from a large corpus of text based on co-occurrence frequency. Thus this approach *implicitly* classifies animal names according to a number of empirically determined dimensions. In contrast, Troyer's et al. (1997) approach relies on several *explicit* semantic dimensions. The choice of a small number of dimensions for manual analysis is quite reasonable because higher dimensionality would prove to be exponentially more difficult for manual human assessments. However, a classification along this practical but small number of dimensions may not be able to capture the variety of possible ways in which people categorize concrete objects such as animals. However, one potential disadvantage of the LSA-based approach, as compared to Troyer's approach, is that the dimensions of the semantic space resulting from LSA computation are not labeled and may not be as readily interpretable as the explicit manually determined dimensions.

4.2 Behavioral Measurements Across Diagnostic Groups

The results with respect to the traditional SVF scores show that they are strongly associated with diagnostic group membership (lowest SVF score in the AD dementia group, highest in the CN group, with MCI in the middle). Standard SVF scores are also strongly associated with STMS, a global cognitive measures of mental status. These results are consistent with prior work showing sensitivity of the SVF score to AD dementia (Henry *et al.*, 2004) and confirm that the SVF performance in the sample used in this study is modulated by cognitive impairment affecting generative verbal fluency. Prior work on qualitative assessments of clustering performance on SVF tests is inconsistent with respect to the effects of MCI and AD dementia on cluster size. Some of the studies found significant differences between healthy controls (larger clusters) and patients with AD (Fagundo et al., 2008; Price et al., 2012; Troster et al., 1998; Troyer et al., 1998), while others failed to find such differences (Epker, Lacritz, & Munro Cullum, 1999; Raoux et al., 2008). The findings of the current study are consistent with the latter - we also did not find a significant difference between diagnostic groups on either the MCS or the CuRel measures. As has been noted by Raoux et al. (2008), differences in qualitative methods and sampling may be responsible for these discrepant results, highlighting the need for computerized quantitative clustering analyses. For example, AD dementia patients in previous studies that found a significant difference in MCS were more impaired on the SVF test (SVF score averaged across studies - 7.73 (SD - 4.3)) than AD dementia patients in studies that did not find significant differences (average SVF score - 9.44 (SD - 3.8)). A difference of 2 words at the lower end of the scale can have a much more significant impact on cluster size computation than the same difference at the

higher end of the scale because the very notion of clusters becomes meaningless as the total number of words to be clustered approaches one. The mean SVF score for the AD dementia group in the current study is 11 (SD 4.6), which reflects the fact that our participants are in earlier stages of dementia. Thus our results are more consistent in this respect with studies that involved higher-functioning participants and did not find significant differences in clustering measures.

One of the objectives of this study was to determine if the overall SVF performance can be decomposed into components that may be associated with semantic processing. The fact that there was no significant association between MCS or CuRel scores and STMS, but that a significant association was found with the traditional SVF score supports the hypothesis that semantic measures (MCS and CuRel) tap into semantic-specific mechanisms rather than those involved in global functioning reflected by STMS scores. On the other hand, given that multiple mechanisms including executive function, attention, inhibition, motor, and semantic retrieval and organization systems are involved in producing speech on semantic verbal fluency tasks, it is not surprising that the traditional SVF score is strongly associated with global cognitive functioning.

4.3 Distinct Neural Mechanisms Underlying Semantic Indices

The task free fMRI imaging results show the intrinsic organization of large-scale neural networks in a task-free experimental paradigm from the perspective of the language network. Our results help elucidate the relationship of the networks to the traditional SVF score and other automatically derived semantic indices including MCS and CuRel. These results suggest that the SVF score, MCS and CuRel measures are dependent on different aspects of brain function, because each has an association with a distinct part of the ICN.

Brain regions found to be associated with SVF, MCS and CuRel metrics have been previously reported in the literature to be related to functions consistent with what one would expect these measures to reflect. The association between SVF scores and loss of positive functional connectivity in the left middle temporal gyrus and left basal ganglia may be indicative of impairment in fronto-temporal network interactions, which are coordinated by subcortical structures (Thames et al., 2012). These structures have been shown to support performance on generative semantic fluency (Libon et al., 2009) and to be involved specifically with lexical and semantic search and retrieval processes but not processes involved in overlearned automatic speech (Birn et al., 2010). Our results are also in-line with a previous task free fMRI study that found left middle temporal connectivity with various frontal and parietal brain regions was predictive of semantic processing in healthy individuals (Wei et al., 2012). The association between SVF scores and gain of positive functional connectivity of the medial prefrontal cortex, which constitutes a part of the aDMN, may index decreased ability to attenuate the activity of the default mode network during the SVF task resulting in interference and consequently poorer performance on the task. This finding is consistent with the previously observed association between increases in connectivity in the aDMN during normal aging, declining STMS scores, and progression to AD dementia (Jones et al., 2011).

Similarly to SVF scores, the association between the MCS measure and gain of positive functional connectivity of the left middle temporal region with the rest of the language network is consistent with prior reports showing this area to be critical for supporting semantic language processing (Hodges & Patterson, 1995; Mummery et al., 1999; Turken & Dronkers, 2011). The loss of negative functional connectivity in the occipital temporal (fusiform gyrus) and superior parietal regions implies dysfunction in those brain regions, in that they are unable to form the appropriate relationship to the rest of the brain during language network formation, which is associated with suboptimal behavioral performance on aspects of the SVF test dependent on the integrity of those brain regions (e.g., smaller semantic clusters). A likely explanation for this finding has to do with the involvement of these areas in working memory (superior parietal region) (Koenigs, Barbey, Postle, & Grafman, 2009) and within-category identification and processing of dominant and distinctive semantic features (fusiform gyrus) (Mechelli, Sartori, Orlandi, & Price, 2006). The finding of the association with fusiform gyrus is also consistent with other fMRI studies of semantic relatedness demonstrating involvement of this area in processing of semantically related, unrelated, and identical pairs of words referring to concrete objects (Wheatley et al., 2005) and processing of object's visual attributes and structural identity (Bruffaerts et al., 2013). Our findings are also consistent with the crucial role that the fusiform gyrus plays in semantic processing involved in the SVF task by being a critical part of the brain's semantic network, with hypometabolism in this area highly predictive of SVF scores in a population with semantic dementia (Mion et al., 2010). A comparison of the brain areas associated with LSA-based MCS and manually computed mMCS (data not shown), confirmed that, consistent with the moderate correlation between the measures, they were associated with similar but not identical regions. Both were associated with loss of negative functional connectivity in the superior parietal regions and gain of positive connectivity in the temporal areas. The involvement of temporal areas was more bilateral for mMCS than for MCS and there was no association of mMCS with the fusiform gyrus. This may possibly be due to the fact that Troyer et al.'s guidelines for computing clusters are based on functional and taxonomic distinctions (e.g., Living Environment, Human Use, Zoological Categories) and thus does not include groupings of animals based on visual structural characteristics.

Higher CuRel scores were associated with loss of positive functional connectivity of the left fronto-parietal regions suggests possible impairment of mechanisms such as working memory and attention that underlie episodic memory retrieval in individuals with higher CuRel scores. This is consistent with prior work showing that the strength of connectivity in the left fronto-parietal network plays an important role in retrieval of previously studied information (Iidaka, Matsumoto, Nogawa, Yamamoto, & Sadato, 2006). The association of CuRel measure with loss of negative functional connectivity of the ventral Default Mode Network (vDMN) (right precuneus, medial temporal lobe) is also consistent with prior findings of the association between these areas and use of visual imagery during episodic memory recall of concrete and highly imageable words (Cavanna & Trimble, 2006). Similarly to the relationship between the MCS and the loss of negative functional connectivity in the occipital temporal and superior parietal regions, the relationship between CuRel and the loss of negative functional connectivity in vDMN regions implies the amelioration of functional distinction of that area from the rest of the brain during language

network formation. As shown in Figure 4, larger CuRel scores are differentially associated with overall test performance in groups with higher versus lower mental status. This is an interesting finding in light of the fact that both groups show detrimental associations between larger CuRel scores and loss of positive and negative functional connectivity in areas supporting working memory and attention mechanisms, yet only the higher functioning group shows the same detrimental association between larger CuRel values and SVF scores. This finding indicates that the same brain mechanism is at work in both groups and that optimal performance for individuals with higher cognitive function (as measured by STMS) consists largely of producing more semantically diverse responses. These results may also indicate the presence of a floor effect in individuals with lower SVF scores that is also evident to some extent in the association between gain of positive functional connectivity in the medial prefrontal cortex (Figure 4a).

In summary, our findings suggest that the computerized semantic indices capture components of the verbal behavior produced on the SVF task that are different from the traditional SVF score. Based on the findings of the current and multiple previous studies (Baldo, Schwartz, Wilkins, & Dronkers, 2006; Mummery, Patterson, Hodges, & Wise, 1996; Tupak et al., 2012), the traditional SVF score indexes the function of frontal and temporal brain areas thought to be involved in executive control and semantic memory necessary for adequate performance on generative verbal fluency tasks. The current study further suggests that the traditional SVF score may not be sufficient to reflect the complexity of brain mechanisms involved in this task. Computational linguistic analysis of responses to this task provide a more complete, albeit most likely not exhaustive account. The MCS measure appears to be associated with the function of parietal and occipital, in addition to temporal areas, shown to be involved in working memory and semantic feature processing. This makes sense in light of the fact that MCS reflects the size of semantic clusters whose production is dependent not only on the availability of conceptual information but also on the ability to use a strategy for their retrieval. A successful strategy is likely to depend on the ability to identify relevant semantic features to produce within-category exemplars, and the ability to keep information online to avoid repetition. The CuRel measure is different from MCS in that it captures the degree of overall semantic relatedness among all concepts produced on the SVF task. Thus it can be thought of as a measure of diversity of the semantic space (i.e., animals in this case) that may be sensitive to impairment in areas involved in memory of previously learned information. It makes sense that individuals with more intact memory would be able to produce more semantically diverse responses – e.g. produce smaller size clusters but a greater number of them due to availability of more previously learned exemplars of a given category. The fact that SVF, MCS and CuRel are uncorrelated with each other and are associated with network connectivity in distinct brain areas suggests that these indices may serve as independent and complementary variables in the investigation of the relationship between brain function and verbal behavior.

4.4 Other Approaches to Clustering Analysis of SVF Responses

Several approaches inspired by graph-theoretic network analysis have been proposed for the analysis of responses on SVF tests (Chan et al., 1993, 2001; Goni et al., 2011; Lerner, Ogrocki, & Thomas, 2009). These approaches estimate the strength of semantic relatedness

between words based on their proximity in individual SVF responses averaged across a large number of subjects to create lexical-semantic networks and describe their properties using similar graph-theoretic methods used in brain network analyses. These approaches use neurodegenerative disease status to group individual responses and to generate and compare the emerging lexical-semantic networks across groups. The underlying assumption in these approaches is that words produced in close proximity to each other on an SVF test are more semantically related than words that are further apart. This assumption may not always be true, particularly for people with neurodegenerative disease. Also, SVF responses typically contain more switches than clusters (mean cluster size is typically between 1 and 2 words). This may present a significant source of “noise” for the analyses that rely on word proximity alone and require a large number of samples for maximum separation between co-occurrences observed and those that are expected by chance. The advantage of the LSA-based approach presented in this paper is that it does not assume that proximity of words in SVF responses defined their semantic relatedness or similarity. Instead, this approach relies on a source of information external to and independent of the SVF testing responses provided by a particular group and thus can be applied to analyze individual responses.

In our own prior work (Pakhomov, Hemmy, & Lim, 2012), we investigated the use of an alternative ontology-based method for calculating the strength of semantic association between pairs of words that was based on the distance between words in WordNet, a large manually constructed lexical database that contains mostly hierarchical relations. The corpus-based approach presented in the current study is complementary to the ontology-based approach. The hierarchical ontology-based approaches tend to measure the degree of similarity between concepts (e.g. cat vs. panther), while the corpus-based approaches tend to also measure the degree of semantic relatedness (e.g., cat vs. mouse). Furthermore, corpus-based approaches tend to be easier and less resource intensive to maintain and update over time; however, the ontology-based approaches tend to be more tractable and explicit (i.e., easier to determine the factors contributing to the various scores). In future work, we plan to experiment with hybrid approaches to computing SVF semantic indices that would leverage both the ontology-based as well as the corpus-based approaches to characterizing strategies used in response to the SVF task.

4.5 Limitations

For the current study, we started with the simplest and most straightforward approach to using LSA to represent word meanings in which each animal was represented as a document from a collection and each document was represented as an unordered list of words - “bag of words”. More sophisticated approaches may include targeting specific words in a predefined context such as a sentence or other syntactic units (Baroni & Lenci, 2010; Devereux, Pilkington, Poibeau, & Korhonen, 2009; Padó & Lapata, 2007), and combining text corpus co-occurrence data with word association norms (Andrews, Vigliocco, & Vinson, 2009; Durda, Buchanan, & Caron, 2009). A number of non-corpus based and hybrid alternatives for computing semantic similarity and relatedness have also been developed, including methods that rely on calculating path length between concepts in a hierarchy such as WordNet (Fellbaum, 1998; G. A. Miller, 1995), as well as methods based on overlap in word definitions in addition to corpus statistics (Jiang & Conrath, 1997; Lesk, 1986;

Patwardhan & Pedersen, 2006; Pedersen, Patwardhan, & Michelizzi, 2004; Resnik, 1999; Resnik, 1995).

6. Conclusion

A number of distinct aspects of verbal behavior exhibited on responses to semantic verbal fluency tests can be described and quantified using automated computational linguistic approaches. In the present study, we examined two different techniques that use the strength of semantic relatedness information to quantify clustering behavior (MCS) as well as the degree of semantic diversity present in the test response (CuRel). These two approaches appear to be supported by neural mechanisms not only distinct from each other, but also from those that support the overall performance on this task (SVF score). These findings have important implications for future studies focused on neural correlates of verbal behavior on generative verbal fluency tasks. Using these objective and relatively easily quantifiable measurements will enable a sensitive, systematic and fine-grained examination of the brain mechanisms involved in lexical access and retrieval. Future work may focus on introducing timing information into the examination of clustering and switching behavior on this task by using audio-recorded samples, as well task-related fMRI imaging to complement the findings obtained in the task-free fMRI paradigm.

Acknowledgements

GRANT SUPPORT: This work was supported by NIH grants R01 LM009623 (PI – Pakhomov), P50 AG016574 (PI – Ron Petersen), U01 AG006786 (PI – Ron Petersen), R01 AG041851 (PIs – David Knopman and Cliff Jack), R01 AG011378 (PI – Cliff Jack), and the Robert H. and Clarice Smith and Abigail Van Buren Alzheimer’s Disease Research Program of the Mayo Foundation. We are also deeply grateful to the participants in this study, their families, and caregivers.

References

- American Psychiatric Association. DSM-IV: Diagnostic and Statistical Manual of Mental Disorders. 4th ed.. Washington, DC: American Psychiatric Association; 1994.
- Andrews M, Vigliocco G, Vinson D. Integrating experiential and distributional data to learn semantic representations. *Psychological Review*. 2009; 116(3):463–498. [PubMed: 19618982]
- Baldo JV, Schwartz S, Wilkins D, Dronkers NF. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *J Int Neuropsychol Soc*. 2006; 12:896–900. [PubMed: 17064451]
- Baroni M, Lenci A. Distributional Memory: A General Framework for Corpus-Based Semantics. *Computational Linguistics*. 2010; 36(4):673–721.
- Baroni M, Murphy B, Barbu E, Poesio M. Strudel: A Corpus-Based Semantic Model Based on Properties and Types. *Cognitive Science*. 2010; 34(2):222–254. [PubMed: 21564211]
- Beckmann C, Mackay C, Filippini N, Smith S. Group comparison of resting-state fMRI data using multi-subject ICA and dual regression. *NeuroImage*. 2009; 47:S148.
- Birn RM, Kenworthy L, Case L, Caravella R, Jones TB, Bandettini PA, Martin A. Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. *Neuroimage*. 2010; 49:1099–1107. [PubMed: 19632335]
- Blei D, Griffiths T, Jordan M, Lafferty J. Latent Dirichlet Allocation. *Journal of Machine Learning Research*. 2003; 3:993–1022.
- Bruffaerts R, Dupont P, De Grauwe S, Peeters R, De Deyne S, Storms G, Vandenberghe R. Right fusiform response patterns reflect visual object identity rather than semantic similarity. *Neuroimage*. 2013; 83:87–97. [PubMed: 23811413]

- Cabeza R, Nyberg L. Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci*. 2000; 12:1–47. [PubMed: 10769304]
- Calhoun VD, Adali T, Pearlson GD, Pekar JJ. A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp*. 2001; 14:140–151. [PubMed: 11559959]
- Caramazza A, Mahon BZ. The organisation of conceptual knowledge in the brain: The future’s past and some future directions. *Cogn Neuropsychol*. 2006; 23:13–38. doi:741532705 [pii] 10.1080/02643290542000021. [PubMed: 21049320]
- Cavanna AE, Trimble MR. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*. 2006; 129:564–583. [PubMed: 16399806]
- Chan AS, Butters N, Salmon DP, McGuire KA. Dimensionality and clustering in the semantic network of patients with Alzheimer’s disease. *Psychol Aging*. 1993; 8:411–419. [PubMed: 8216961]
- Chan AS, Salmon DP, De La Pena J. Abnormal semantic network for “animals” but not “tools” in patients with Alzheimer’s disease. *Cortex*. 2001; 37:197–217. [PubMed: 11394721]
- Chao-Gan Y, Yu-Feng Z. DPARSF: A MATLAB Toolbox for “Pipeline” Data Analysis of Resting-State fMRI. *Front Syst Neurosci*. 2010; 4:13. [PubMed: 20577591]
- Devereux B, Pilkington N, Poibeau T, Korhonen A. Towards Unrestricted, Large-Scale Acquisition of Feature-Based Conceptual Representations from Corpus Data. *Research on Language and Computation*. 2009; 7(2–4):137–170.
- Durda K, Buchanan L, Caron R. Grounding co-occurrence: Identifying features in a lexical co-occurrence model of semantic memory. *Behavior Research Methods*. 2009; 41(4):1210–1223. [PubMed: 19897830]
- Epker MO, Lacritz LH, Munro Cullum C. Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *J Clin Exp Neuropsychol*. 1999; 21:425–434. [PubMed: 10550803]
- Erhardt EB, Rachakonda S, Bedrick EJ, Allen EA, Adali T, Calhoun VD. Comparison of multi-subject ICA methods for analysis of fMRI data. *Human Brain Mapping*. 2011; 32(12):2075–2095. [PubMed: 21162045]
- Estes WK. Learning theory and intelligence. *Am Psychol*. 1974; 29:740–749.
- Fagundo AB, Lopez S, Romero M, Guarch J, Marcos T, Salamero M. Clustering and switching in semantic fluency: predictors of the development of Alzheimer’s disease. *Int J Geriatr Psychiatry*. 2008; 23:1007–1013. [PubMed: 18416452]
- Fellbaum, C., editor. *Word Net: an electronic lexical database*. Cambridge, Mass: MIT Press; 1998.
- Fox MD, Zhang D, Snyder AZ, Raichle ME. The global signal and observed anticorrelated resting state brain networks. *J Neurophysiol*. 2009; 101:3270–3283. [PubMed: 19339462]
- Goni J, Arrondo G, Sepulcre J, Martincorena I, Velez de Mendizabal N, Corominas-Murtra B, Villoslada P. The semantic organization of the animal category: evidence from semantic verbal fluency and network theory. *Cognitive Processing*. 2011; 12:183–196. [PubMed: 20938799]
- Grossman M. Frontotemporal dementia: a review. *J Int Neuropsychol Soc*. 2002; 8:566–583. [PubMed: 12030310]
- Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer’s type: a meta-analysis. *Neuropsychologia*. 2004; 42:1212–1222. doi:10.1016/j.neuropsychologia.2004.02.001 S0028393204000296 [pii]. [PubMed: 15178173]
- Hodges JR, Davies RR, Xuereb JH, Casey B, Broe M, Bak TH, Halliday GM. Clinicopathological Correlates in Frontotemporal Dementia. *Ann Neurol*. 2004; 56:399–406. [PubMed: 15349867]
- Hodges JR, Patterson H. Is semantic memory consistently impaired early in the course of Alzheimer’s disease. Neuroanatomical and diagnostic implications. *Neuropsychologia*. 1995; 33:441–459. [PubMed: 7617154]
- Iidaka T, Matsumoto A, Nogawa J, Yamamoto Y, Sadato N. Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cereb Cortex*. 2006; 16:1349–1360. [PubMed: 16861334]
- Jiang, J.; Conrath, D. Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy; Proceedings of the International Conference on Research in Computational Linguistics; 1997. p. 19-33.

- Jones DT, Machulda MM, Vemuri P, McDade EM, Zeng G, Senjem ML, Jack CR. Age-related changes in the default mode network are more advanced in Alzheimer disease. *Neurology*. 2011; 77:1524–1531. [PubMed: 21975202]
- Jones DT, Vemuri P, Murphy MC, Gunter JL, Senjem ML, Machulda MM, Jack CR. Non-Stationarity in the “Resting Brain’s” Modular Architecture. *PLoS One*. 2012; 7:e39731. [PubMed: 22761880]
- Kanerva P. Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cognitive Computation*. 2009; 1(2):139–159.
- Keller CJ, Bickel S, Entz L, Ulbert I, Milham MP, Kelly C, Mehta AD. Intrinsic functional architecture predicts electrically evoked responses in the human brain. *Proc Natl Acad Sci U S A*. 2011; 108:10308–10313. [PubMed: 21636787]
- Knopman DS, Kramer JH, Boeve BF, Caselli RJ, Graff-Radford NR, Mendez MF, Mercaldo N. Development of methodology for conducting clinical trials in frontotemporal lobar degeneration. *Brain*. 2008; 131:2957–2968. [PubMed: 18829698]
- Koenigs M, Barbey AK, Postle BR, Grafman J. Superior parietal cortex is critical for the manipulation of information in working memory. *J Neurosci*. 2009; 29:14980–14986. [PubMed: 19940193]
- Kokmen E, Smith GE, Petersen RC, Tangalos E, Ivnik RC. The short test of mental status. Correlations with standardized psychometric testing. *Arch Neurol*. 1991; 48:725–728. [PubMed: 1859300]
- Laine, M. Correlates of word fluency performance. In: Koivuselka-Sallinen, P.; Sarajarvi, L., editors. *Studies in Languages*. Vol. 12. Joensuu, Finland: University of Joensuu; 1988.
- Landauer, TK. *Handbook of latent semantic analysis*. Mahwah, N.J.: Lawrence Erlbaum Associates; 2006.
- Landauer TK, Dumais ST. A solution to Plato’s problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychol Rev*. 1997; 104:211–240.
- Lerner AJ, Ogrocki PK, Thomas PJ. Network graph analysis of category fluency testing. *Cogn Behav Neurol*. 2009; 22:45–52. doi:10.1097/WNN.0b013e318192ccaf00146965-200903000-00005 [pii]. [PubMed: 19372770]
- Lesk, M. Automatic Sense Disambiguation Using Machine Readable Dictionaries: How to Tell a Pine Cone from an Ice Cream Cone; Proceedings of the 5th Annual International Conference on System Documentation; 1986. p. 24-26.
- Levelt, WJ. *Speaking: from intention to articulation*. Cambridge, MA: MIT Press; 1989.
- Lezak, MD. *Neuropsychological Assessment*. 4th ed.. Oxford, England: Oxford University Press; 2004.
- Libon DJ, McMillan C, Gunawardena D, Powers C, Massimo L, Khan A, Grossman M. Neurocognitive contributions to verbal fluency deficits in frontotemporal lobar degeneration. *Neurology*. 2009; 73:535–542. [PubMed: 19687454]
- Mahon BZ, Caramazza A. Concepts and categories: a cognitive neuropsychological perspective. *Annu Rev Psychol*. 2009; 60:27–51. [PubMed: 18767921]
- McClelland JL, Rogers TT. The parallel distributed processing approach to semantic cognition. *Nat Rev Neurosci*. 2003; 4:310–322. [PubMed: 12671647]
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer’s disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease. *Neurology*. 1984; 34:939–944. [PubMed: 6610841]
- Mechelli A, Sartori G, Orlandi P, Price CJ. Semantic relevance explains category effects in medial fusiform gyri. *Neuroimage*. 2006; 30:992–1002. [PubMed: 16343950]
- Meinzer M, Flaisch T, Wilser L, Eulitz C, Rockstroh B, Conway T, Crosson B. Neural signatures of semantic and phonemic fluency in young and old adults. *J Cogn Neurosci*. 2009; 21:2007–2018. [PubMed: 19296728]
- Miller GA. WordNet: a lexical database for English. *Communications of the ACM*. 1995; 38(11):39–41.
- Miller G, Fellbaum C. Semantic Networks of English. *Cognition*. 1991; 41:197–229. [PubMed: 1790654]

- Mion M, Patterson K, Acosta-Cabronero J, Pengas G, Izquierdo-Garcia D, Hong YT, Nestor PJ. What the left and right anterior fusiform gyri tell us about semantic memory. *Brain*. 2010; 133:3256–3268. [PubMed: 20952377]
- Mitchell TM, Shinkareva SV, Carlson A, Chang K-M, Malave VL, Mason RA, Just MA. Predicting human brain activity associated with the meanings of nouns. *Science (New York, N Y)*. 2008; 320(5880):1191–1195.
- Mummery CJ, Patterson K, Hodges JR, Wise RJ. Generating “tiger” as an animal name or a word beginning with T: differences in brain activation. *Proc Biol Sci*. 1996; 263:989–995. [PubMed: 8805836]
- Mummery CJ, Patterson K, Wise RJ, Vandenberghe R, Price CJ, Hodges JR. Disrupted temporal lobe connections in semantic dementia. *Brain*. 1999; 122(Pt 1):61–73. [PubMed: 10050895]
- Ober BA, Dronkers NF, Koss E, Delis DC, Friedland RP. Retrieval from semantic memory in Alzheimer-type dementia. *J Clin Exp Neuropsychol*. 1986; 8:75–92. [PubMed: 3944246]
- Padó S, Lapata M. Dependency-Based Construction of Semantic Space Models. *Computational Linguistics*. 2007; 33(2):161–199.
- Pakhomov S, Hemmy L. A computational linguistic measure of clustering behavior on semantic verbal fluency task predicts risk of future dementia in the Nun Study. *Cortex*. 2013 (in press), (e-pub ahead of print).
- Pakhomov S, Hemmy LS, Lim KO. Automated semantic indices related to cognitive function and rate of cognitive decline. *Neuropsychologia*. 2012; 50:2165–2175. [PubMed: 22659109]
- Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci*. 2007; 8:976–987. doi:nrn2277 [pii] 10.1038/nrn2277. [PubMed: 18026167]
- Patwardhan S, Pedersen T. Using WordNet-based Context Vectors to Estimate the Semantic Relatedness of Concepts. In *Proceedings of the EACL 2006 Workshop on Making Sense of Sense: Bringing Computational Linguistics and Psycholinguistics Together*. 2006:1–8.
- Pedersen T, Pakhomov SV, Patwardhan S, Chute CG. Measures of semantic similarity and relatedness in the biomedical domain. *J Biomed Inform*. 2007; 40:288–299. doi:S1532-0464(06)00064-5 [pii] 10.1016/j.jbi.2006.06.004. [PubMed: 16875881]
- Pedersen, T.; Patwardhan, S.; Michelizzi, J. WordNet::Similarity - Measuring the Relatedness of Concepts; Proceedings of Fifth Annual Meeting of the North American Chapter of the Association for Computational Linguistics; 2004. p. 38-41.
- Pereira F, Botvinick M, Detre G. Using Wikipedia to learn semantic feature representations of concrete concepts in neuroimaging experiments. *Artificial Intelligence*. 2013; 194:240–252. [PubMed: 23243317]
- Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med*. 2004; 256:183–194. [PubMed: 15324362]
- Price SE, Kinsella GJ, Ong B, Storey E, Mullaly E, Phillips M, Perre D. Semantic verbal fluency strategies in amnesic mild cognitive impairment. *Neuropsychology*. 2012; 26:490–497. [PubMed: 22746308]
- Prince SE, Tsukiura T, Cabeza R. Distinguishing the neural correlates of episodic memory encoding and semantic memory retrieval. *Psychological Science*. 2007; 18:144–151. [PubMed: 17425535]
- Pylkkanen L, Llinas R, Murphy GL. The representation of polysemy: MEG evidence. *J Cogn Neurosci*. 2006; 18:97–109. [PubMed: 16417686]
- Quesada, J. Creating Your Own LSA Spaces. In: Landauer, T.; McNamara, D.; Dennis, S.; Kintsch, W., editors. *Handbook of Latent Semantic Analysis*. New York: Taylor and Francis Group; 2011. p. 71-89.
- Rada, R.; Mili, F.; Bicknell, E.; Blettner, AM. Development and application of a metric on semantic nets. Vol. 19. Presented at the IEEE Transaction on Systems, Man, and Cybernetics, IEEE; 1989. p. 17-30.
- Raoux N, Amieva H, Le Goff M, Auriacombe S, Carcaillon L, Letenneur L, Dartigues JF. Clustering and switching processes in semantic verbal fluency in the course of Alzheimer’s disease subjects: results from the PAQUID longitudinal study. *Cortex*. 2008; 44:1188–1196. doi:S0010-9452(07)00158-X [pii] 10.1016/j.cortex.2007.08.019. [PubMed: 18761132]

- Resnik, P. Using Information Content to Evaluate Semantic Similarity in a Taxonomy; Proceedings of the 14th International Joint Conference on Artificial Intelligence; 1995. p. 448-453.
- Resnik P. Semantic similarity in a taxonomy: an information-based measure and its application to problems of ambiguity in natural language. *J of Artif Intell Res.* 1999; 11:95–130.
- Rich JB, Troyer AK, Bylsma FW, Brandt J. Longitudinal analysis of phonemic clustering and switching during word-list generation in Huntington's disease. *Neuropsychology.* 1999; 13:525–531. [PubMed: 10527060]
- Rosen WG. Verbal fluency in aging and dementia. *J Clin Neuropsychol.* 1980; 2:135–146.
- Salmon DP, Butters N, Chan AS. The deterioration of semantic memory in Alzheimer's disease. *Can J Exp Psychol.* 1999; 53:108–117. [PubMed: 10389493]
- Song XW, Dong ZY, Long XY, Li SF, Zuo XN, Zhu CZ, Zang YF. REST: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS One.* 2011; 6:e25031. [PubMed: 21949842]
- Stam CJ, de Haan W, Daffertshofer A, Jones BF, Manshanden I, van Cappellen van Walsum AM, Scheltens P. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain.* 2009; 132:213–224. doi:aw262 [pii] 10.1093/brain/aw262. [PubMed: 18952674]
- Stam CJ, Jones BF, Nolte G, Breakspear M, Scheltens P. Small-world networks and functional connectivity in Alzheimer's disease. *Cereb Cortex.* 2007; 17:92–99. [PubMed: 16452642]
- Supekar K, Menon V, Rubin D, Musen M, Greicius MD. Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Computational Biology.* 2008; 4:e1000100. [PubMed: 18584043]
- Tang-Wai DF, Knopman DS, Geda YE, Edland SD, Smith GE, Ivnik RJ, Petersen RC. Comparison of the short test of mental status and the mini-mental state examination in mild cognitive impairment. *Arch Neurol.* 2003; 60:1777–1781. [PubMed: 14676056]
- Thames AD, Foley JM, Wright MJ, Panos SE, Ettenhofer M, Ramezani A, Hinkin CH. Basal ganglia structures differentially contribute to verbal fluency: evidence from Human Immunodeficiency Virus (HIV)-infected adults. *Neuropsychologia.* 2012; 50:390–395. [PubMed: 22223078]
- Troster AI, Fields JA, Testa JA, Paul RH, Blanco CR, Hames KA, Beatty WW. Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia.* 1998; 36:295–304. [PubMed: 9665640]
- Troyer AK. Normative data for clustering and switching on verbal fluency tasks. *J Clin Exp Neuropsychol.* 2000; 22:370–378. [PubMed: 10855044]
- Troyer AK, Moscovitch M, Winocur G. Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology.* 1997; 11:138–146. [PubMed: 9055277]
- Troyer AK, Moscovitch M, Winocur G, Leach L, Freedman M. Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *J Int Neuropsychol Soc.* 1998; 4:137–143. [PubMed: 9529823]
- Tulving, E. Episodic and semantic memory. In: Tulving, E.; Donaldson, W., editors. *Organization of Memory.* Vol. 381. New York: Academic Press; 1972. p. 381-402.
- Tulving E, Kapur S, Craik FI, Moscovitch M, Houle S. Hemispheric encoding/retrieval asymmetry in episodic memory: positron emission tomography findings. *Proc Natl Acad Sci U S A.* 1994; 91:2016–2020. [PubMed: 8134342]
- Tupak SV, Badewien M, Dresler T, Hahn T, Ernst LH, Herrmann MJ, Ehlis AC. Differential prefrontal and frontotemporal oxygenation patterns during phonemic and semantic verbal fluency. *Neuropsychologia.* 2012; 50:1565–1569. [PubMed: 22426205]
- Turken, Dronkers NF. The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Front Syst Neurosci.* 2011; 5:1. [PubMed: 21347218]
- Tverski B, Hemenway K. Objects, parts and categories. *J Exp Psychol.* 1984; 113:169–197.
- Weber M, Thompson-Schill SL, Osherson D, Haxby J, Parsons L. Predicting judged similarity of natural categories from their neural representations. *Neuropsychologia.* 2009; 47:859–868. doi:S0028-3932(08)00504-6 [pii] 10.1016/j.neuropsychologia.2008.12.029. [PubMed: 19162048]

- Wei T, Liang X, He Y, Zang Y, Han Z, Caramazza A, Bi Y. Predicting Conceptual Processing Capacity from Spontaneous Neuronal Activity of the Left Middle Temporal Gyrus. *Journal of Neuroscience*. 2012; 32(2):481–489. [PubMed: 22238084]
- Weissenbacher A, Kasess C, Gerstl F, Lanzenberger R, Moser E, Windischberger C. Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies. *Neuroimage*. 2009; 47:1408–1416. [PubMed: 19442749]
- Wheatley T, Weisberg J, Beauchamp MS, Martin A. Automatic priming of semantically related words reduces activity in the fusiform gyrus. *J Cogn Neurosci*. 2005; 17:1871–1885. [PubMed: 16356325]

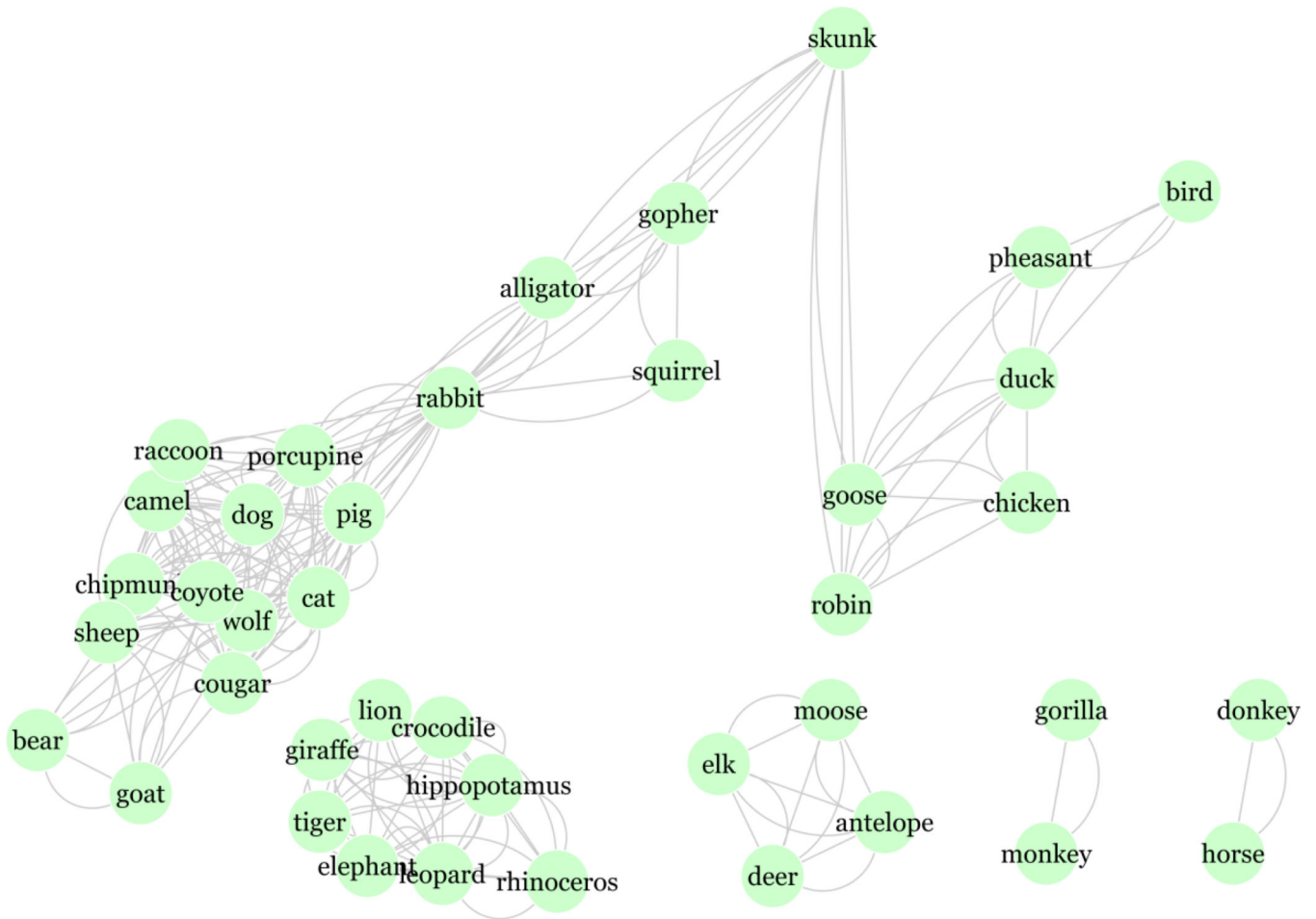


Figure 1. Network of top 50 most frequent animal names found in the study participants' responses. This network was constructed based on applying a threshold (0.9) to pairwise semantic relatedness values computed from the dimensionality-reduced LSA matrix.

Subject characteristics:
male, 77 y.o., diag: MCI, STMS: 36, CDR: 0.5

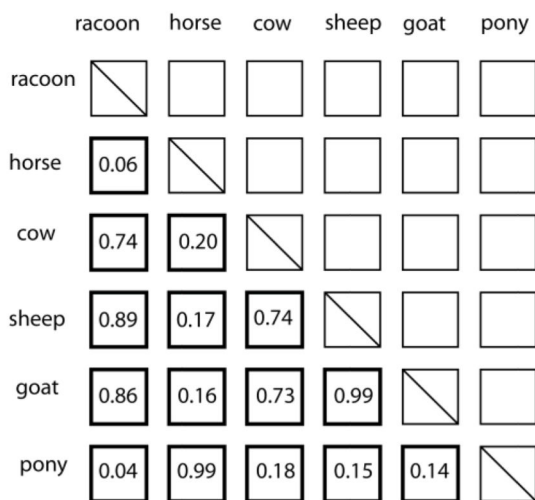
Subject characteristics:
male, 80 y.o., diag: AD, STMS: 30, CDR: 0.5

SVF sample 1:

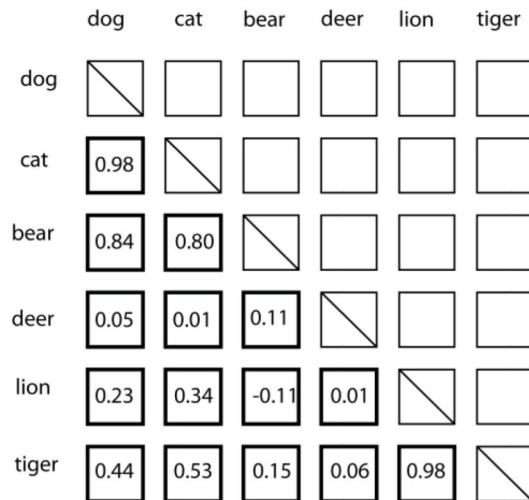
	raccoon	horse	cow	sheep	goat	pony
cluster number	1	2	3	4	5	
cluster size	0	0	0	1	0	

SVF sample 2:

	dog	cat	bear	deer	lion	tiger
cluster number	1	2	3		4	
cluster size	1	0	0		1	



cumulative relatedness (CuRel = 0.47)
mean cluster size (MCS: $0+0+0+1+0 / 5 = 0.20$)
number of clusters = 5



cumulative relatedness (CuRel = 0.36)
mean cluster size (MCS: $1+0+0+1 / 4 = 0.50$)
number of clusters = 4

Figure 2.
Illustration of cumulative semantic relatedness (CuRel) computation.

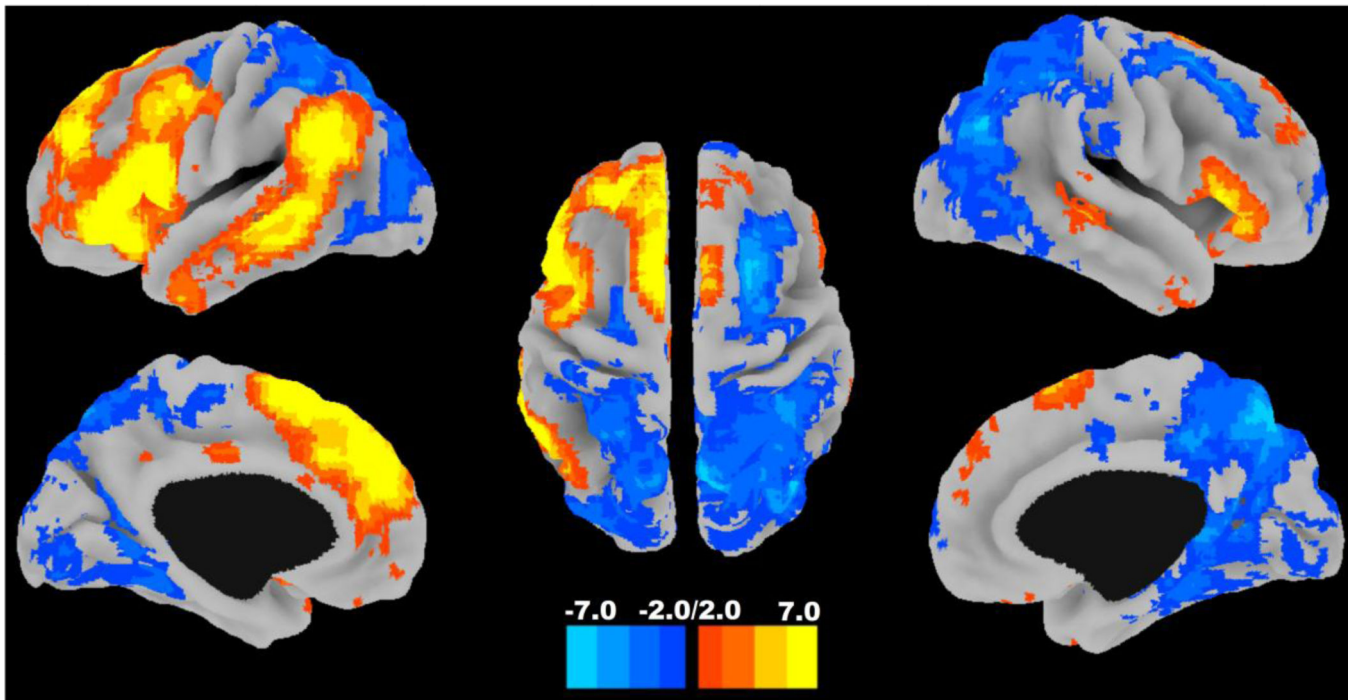


Figure 3.

Group map of back-reconstructed language network. The one-sample t-test of all of the subjects' back-reconstructed language networks is displayed on a surface rendering created using the Caret software package (<http://www.nitrc.org/projects/caret/>). The left hemisphere is presented in the left half of the image. The color bar encodes the t-statistic.

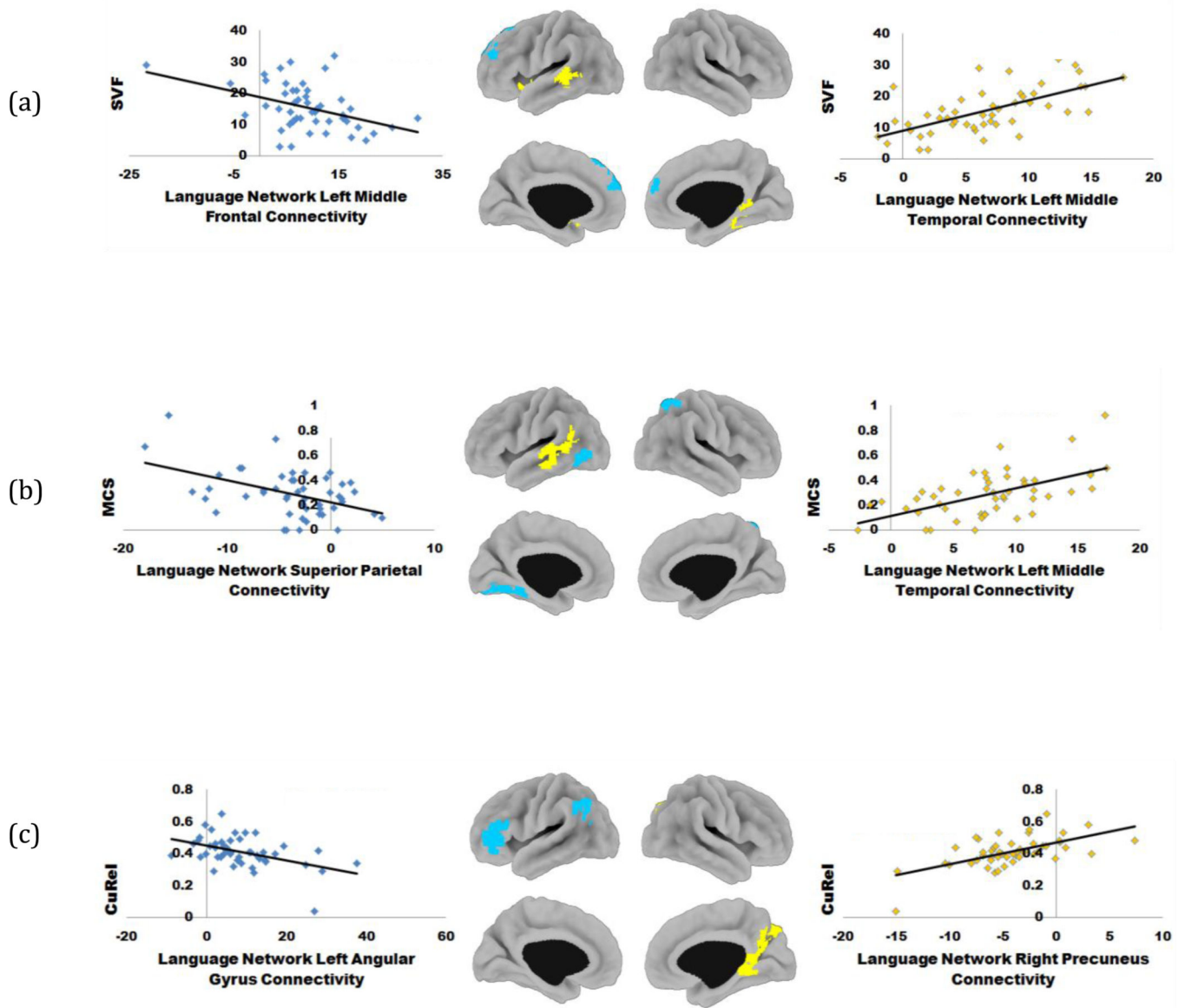


Figure 4. Task free fMRI network connectivity associated with several aspects of performance on SVF tests. (a) – areas of positive (yellow) and negative (blue) correlation between the standard SVF score and language network connectivity; (b) – areas of positive and negative correlation between MCS and language network connectivity; (c) – areas of positive and negative correlation between CuRel and language network connectivity.

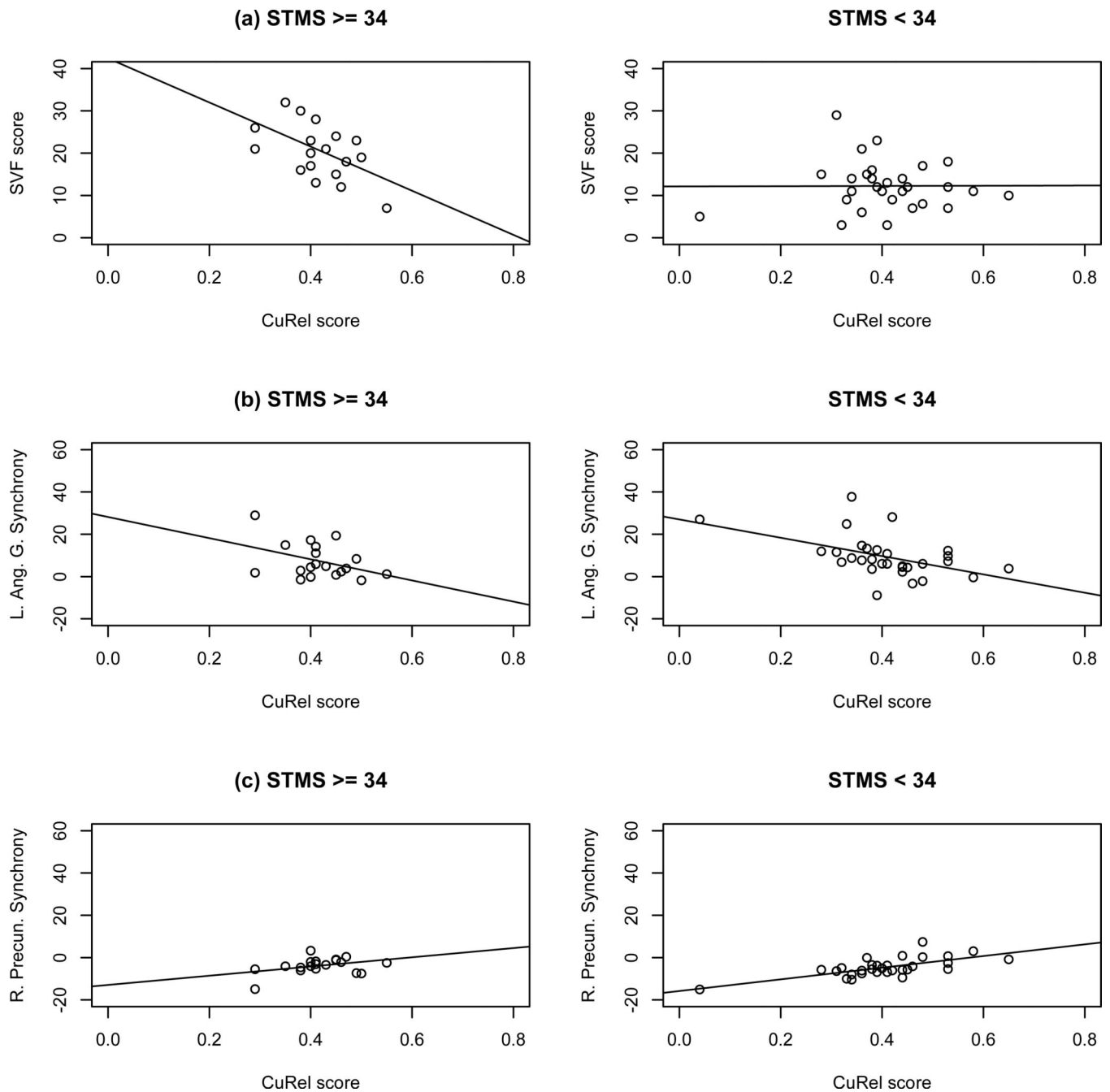


Figure 5. Comparisons of associations between CuRel score and overall performance on the SVF test (a) with loss of positive functional connectivity of the left angular gyrus (b) and loss of negative functional connectivity in the right precuneus (c).

Table 1

Demographic, clinical, and SVF characteristics by diagnostic group.

	CN (N=11)	MCI (N=17)	AD (N=21)	p-value	Tukey HSD
Female (%)	7 (63.6)	5 (29.4)	6 (28.5)	0.11	
Mean Age	55 (13)	69 (12)	69 (11)	0.005	
Mean STMS ¹	36 (1.2)	33 (2.3)	27 (5.6)	<0.001	CN > AD MCI > AD
Mean CDR ²	0 (0)	1.2 (0.78)	3.6 (1.96)	<0.001	CN > AD MCI > AD
Mean SVF ³	23 (6.3)	16 (7.0)	11 (4.6)	<0.001	CN > AD CN > MCI
Mean MCS ⁴	0.26 (0.11)	0.31 (0.24)	0.29 (0.18)	0.79	
Mean CuRel ⁵	0.40 (0.05)	0.40 (0.08)	0.43 (0.12)	0.24	

¹ STMS - short test of mental status score

² CDR - clinical dementia rating

³ SVF - semantic verbal fluency score

⁴ MCS - mean cluster size

⁵ CuRel - raw cumulative semantic relatedness score

Table 2
Brain network connectivity associated with different aspects of performance on SVF test.

Covariate	Cluster #	T-Score	# Voxels	MNI (x,y,z)	ICN	Region	Detrimental Association*
SVF-Pos¹	1	3.98	123	(-57,-27,0)	Language	L. Middle Temporal	Loss of positive connectivity
	2	3.91	202	(-24,15,-12)	Deep Gray	L. Putamen	
	3	3.31	117	(21,-42,6)	vDMN ⁴	R. Precuneus	
SVF-Neg	1	-3.64	212	(-27,51,33)	adDMN ⁵	L. Middle Frontal	Gain of positive connectivity
MCS-Pos²	1	4.2	211	(-51,-57,18)	Language	L. Middle Temporal	Loss of positive connectivity
MCS-Neg	1	-4.08	125	(21,-57,54)	SPL ⁶	SPL	Loss of negative connectivity
	2	-3.63	238	(-48,-63,-12)	Ventral Stream	L. Inf. Occipital	
CuRel-Pos³	1	3.59	303	(15,-54,75)	vDMN	R. Precuneus	Loss of negative connectivity
CuRel-Neg	1	-3.88	157	(-45,-63,45)	Working Memory	L. Angular Gyrus	Loss of positive connectivity
	2	-3.31	206	(-42,39,0)	Language	L. IFG p.Triangularis	

¹ SVF – semantic verbal fluency score (manual)

² MCS – mean cluster size

³ CuRel - cumulative relatedness (automated)

⁴ vDMN-ventral default mode network

⁵ adDMN- anterior dorsal default mode network

⁶ SPL-superior parietal lobule