Changes in functional connectivity related to direct training and generalization effects of a word finding treatment in chronic aphasia

Chaleece W. Sandberg[*], **Jason W. Bohland**, and **Swathi Kiran**
Boston University, Department of Speech and Hearing Sciences, 635 Commonwealth Ave., Boston, MA 02215, USA

Abstract

The neural mechanisms that underlie generalization of treatment-induced improvements in word finding in persons with aphasia (PWA) are currently poorly understood. This study aimed to shed light on changes in functional network connectivity underlying generalization in aphasia. To this end, we used fMRI and graph theoretic analyses to examine changes in functional connectivity after a theoretically-based word-finding treatment in which abstract words were used as training items with the goal of promoting generalization to concrete words. Ten right-handed native English-speaking PWA (7 male, 3 female) ranging in age from 47 to 75 (mean = 59) participated in this study. Direct training effects coincided with increased functional connectivity for regions involved in abstract word processing. Generalization effects coincided with increased functional connectivity for regions involved in concrete word processing. Importantly, similarities between training and generalization effects were noted as were differences between participants who generalized and those who did not.

Keywords

Aphasia; Treatment; fMRI; Functional connectivity; Generalization; Neuroplasticity

1. Introduction

Aphasia is a language deficit which often occurs following a cerebrovascular accident (CVA) of the language dominant hemisphere of the brain. The majority of spontaneous recovery after a CVA occurs within the acute and subacute stages, reaching a plateau during the chronic stage (Cramer, 2008). *Importantly, even in the chronic stage, after spontaneous recovery has ceased, treatment can induce neuroplasticity*. Neuroplasticity related to treatment for aphasia in the chronic phase has been shown in a number of neuroimaging studies in the last decade (e.g., Cornelissen et al., 2003; Fridriksson, Richardson, Fillmore,

& Cai, 2012; Marcotte, Perlbarg, Marrelec, Benali, & Ansaldo, 2013; Meinzer et al., 2008; Rochon et al., 2010; van Hees et al., 2014; Vitali et al., 2010).

An important feature of successful therapy in aphasia is generalization from trained to untrained items, leading to more efficient and cost-effective therapies. While many neuroimaging studies have examined treatment-induced neuroplasticity in aphasia, few have systematically examined and reported neural changes associated with behavioral improvements due to both direct training and generalization effects of treatment (Meinzer et al., 2008; Vitali et al., 2010). The current study examines neural changes related to direct training and generalization effects of a theoretically based treatment for word finding difficulties. We hypothesized that changes to the semantic processing system would emerge as modifications in the functional pathways between a distributed set of brain areas, and that such changes would mirror behavioral changes associated with therapy. Given these hypotheses and the relatively large network of brain regions involved in word finding, we adopted a network-based functional connectivity analysis of fMRI data obtained before and after therapy.

While fMRI activation studies provide valuable insight into local effects of treatment-induced neuroplasticity, functional connectivity analyses can capture changes in task-specific coupling of brain areas related to treatment gains, which may not be evident from activation-based contrasts. Analyses of changes in functional connectivity allow researchers to make inferences not only about the coupling among brain regions while subjects engage certain processes, but also about how that coupling is influenced by changes in the experimental context or over time. Examining changes in functional connectivity is relatively new in aphasia treatment research, with only a few studies published to date (Abutalebi, Rosa, Tettamanti, Green, & Cappa, 2009; Marcotte et al., 2013; Sarasso et al., 2010; van Hees et al., 2014; Vitali et al., 2010). These studies utilize techniques such as Structural Equation Modeling (SEM; Sarasso et al., 2010; Vitali et al., 2010) and Dynamic Causal Modeling (DCM; Abutalebi et al., 2009), and examine both resting state (van Hees et al., 2014) and task-based (Abutalebi et al., 2009; Marcotte et al., 2013; Sarasso et al., 2010; Vitali et al., 2010) functional connectivity.

These studies have shown increased functional connectivity related to direct training, and in one case, to generalization effects. Using SEM, Vitali et al. (2010) found that direct training effects of a phonological cueing treatment in two patients with phonological anomia were evident immediately following treatment, while generalization effects were observed at a six-month follow-up test, when the trained versus untrained conditions showed similar levels of performance. The authors suggest that these patients eventually adopted a strategy for lexical retrieval of untrained items that mirrored their training-induced lexical retrieval strategy.

Based on these recent studies of changes in functional connectivity following treatment in aphasia, we hypothesize that direct training effects will coincide with increased functional connectivity within networks engaged by the trained items (Abutalebi et al., 2009; Sarasso et al., 2010; Vitali et al., 2010) and that generalization effects will coincide with increased functional connectivity in networks engaged by the untrained items (Vitali et al., 2010).

Here we adopt a graph theoretic approach, novel to aphasia treatment research, to test these hypotheses related to treatment-induced changes in functional connectivity in the language system.

Graph theory has been used to characterize complex systems ranging from social networks to molecular interactions and is gaining popularity in cognitive neuroscience as it provides meaningful insights into the organization and dynamics of complex networks such as the brain. Neuroimaging studies that utilize graph theory have begun to identify changes in intrinsic network properties, for example, after motor training (e.g., Sami & Miall, 2013). However, such approaches are currently underutilized in the study of neuroplastic changes related to treatment outcomes, and no study to date has used graph theoretical measures to characterize changes in task-based networks after language treatment. The current study utilizes node degree, which expresses the importance of a region via the quantity of its incoming and outgoing connections within a network, and allows a comparison across participants, while taking into account individual variability of network composition (functional regions of interest included within the network).

One hurdle in examining neural changes related to direct training and generalization effects, especially in a word-finding treatment, is that both the trained and the untrained items are expected to activate the same regions in the semantic network. This is especially true in picture-naming paradigms, which necessarily use concrete words for both trained and untrained items. One way to separate the effects of direct training and generalization is to choose stimuli that have been shown to be dissociable, both behaviorally and neurophysiologically, such as abstract and concrete words. Recent neuroimaging research suggests that abstract and concrete words are preferentially processed in different cortical regions (Binder, Desai, Graves, & Conant, 2009; Wang, Conder, Blitzer, & Shinkareva, 2010). In a meta-analysis of 17 imaging studies, Binder et al. (2009) found preferential activation for concrete concepts in bilateral angular gyrus (AG), left fusiform gyrus (L FFG), left dorsomedial prefrontal cortex (L DMPFC) – which includes portions of superior medial gyrus (SupMed), superior frontal gyrus (SFG), and middle frontal gyrus (MFG) – and left posterior cingulate cortex (L PCC). Preferential activation for abstract concepts was found in left inferior frontal gyrus (L IFG) and left anterior superior temporal sulcus (L aSTS).

In a separate meta-analysis, Wang et al. (2010) examined 19 functional neuroimaging studies of abstract and concrete concept processing, ten of which overlapped with the Binder et al. (2009) study. They found consistent activation in left precuneus (L PCN), L PCC, L FFG, and left parahippocampal gyrus (L PHG) for the [concrete > abstract] contrast and activation in L IFG, L aSTG, left anterior middle temporal gyrus (L aMTG), and left posterior middle temporal gyrus (L pMTG) for the [abstract > concrete] contrast. Similar to the Binder et al. (2009) study, abstract words preferentially activated verbal regions while concrete words preferentially activated nonverbal regions, although the exact regions differed somewhat.

In our own work, we have shown that training abstract words (i.e., *justice*) in a particular context-category (i.e., *courthouse*) promotes generalization to concrete words (i.e., *jury*) in

the same context-category, but not vice versa (Kiran, Sandberg, & Abbott, 2009; Sandberg & Kiran, 2014). This treatment is based on the complexity account of treatment efficacy (Thompson, Shapiro, Kiran, & Sobecks, 2003), and suggests that training semantic features of abstract words – the more complex items – promotes spreading activation to a large distributed network of associated words, including concrete words, whereas training semantic features for concrete words – the less complex items – results in spreading activation to a more confined network of related concrete words. We hypothesize that training abstract words will increase functional connectivity of the abstract network and that generalization to concrete words will increase functional connectivity of the concrete network. However, the underlying neural mechanism of generalization from abstract to concrete words is currently unknown.

Thus, the goal of this study is to shed light on the underlying neural mechanism of generalization from abstract to concrete words by comparing the functional connectivity patterns of abstract and concrete word processing in PWA before and after a successful word generation treatment. This study builds upon the current literature by systematically examining both direct training and generalization effects of treatment, incorporating a functional connectivity analysis, and utilizing graph theoretical measures to characterize changes in functional connectivity related to treatment gains. We also use a sample size of 10 PWA in response to the prevalence of single- or multiple-case studies with fewer than eight PWA, which makes it difficult to examine consistent patterns across PWA. Additionally, we include a no-treatment within-patient control subgroup, as only one study to date has included a no-treatment patient control, albeit not within-patient controls (Rochon et al., 2010), and patient controls, especially within-patient controls are an ideal baseline for measuring treatment gains.

Our specific aims and hypotheses were as follows:

a. *Examine changes in functional connectivity during a no-treatment control period.* We analyzed scan-to-scan changes in functional connectivity for a subgroup of PWA who participated in a no-treatment control period before beginning treatment as a baseline comparison for the effects of treatment. Based on Rochon et al. (2010) and Meinzer et al. (2006), who found scan-to-scan stability in activation for healthy control subjects, we expected scan-to-scan stability in functional connectivity for this subgroup of PWA during the control period, as all PWA were in the chronic stage of recovery.

b. *Examine changes in functional connectivity related to direct training effects of treatment.* We analyzed pre- to post-treatment changes in functional connectivity for the trained items (abstract words) for the group of PWA who improved on the trained abstract words. Based on Abutalebi et al. (2009), Sarasso et al. (2010), and Vitali et al. (2010), we expected direct training effects to coincide with increased functional connectivity in the abstract word network. Based on the neuroimaging literature of abstract and concrete words (e.g., Binder et al., 2009; Wang et al., 2010), we expected that the regions with the largest network-level changes would include L IFG, L STG, and/or L MTG.

c. *Examine changes in functional connectivity related to generalization effects of treatment.* In addition to the training effects for abstract words, we analyzed pre- to post-treatment changes in functional connectivity for the untrained items (concrete words) for the group of PWA who improved on the trained abstract words and generalized to the untrained concrete words. Based on Vitali et al. (2010), we expected generalization effects of treatment to coincide with increased functional connectivity in the concrete word network. Based on the neuroimaging literature of abstract and concrete words (e.g., Binder et al., 2009; Wang et al., 2010), we expected that the regions with the largest network-level changes would include L AG, R AG, L FFG, L PHG, L DMPFC (i.e., L SupMed, L SFG, and/or L MFG), L PCN, and/or L PCC (see Table 1 for a list of region abbreviations that will be used throughout this manuscript).

2. Method

2.1. Participants

Ten right-handed native English-speaking persons with aphasia (7 male, 3 female) ranging in age from 47 to 75 (mean = 59) participated in this study. All participants experienced a CVA in the distribution of the left middle cerebral artery (see Supplementary Fig. 1) and were in the chronic stage of recovery as evidenced by time post-onset of at least six months. All participants had at least a high school education. All patients received generative word-finding therapy as well as pre- and post-treatment fMRI scans. Three patients additionally participated in a 'no treatment' set of scans as a control for changes in activation patterns unrelated to the therapy (see Fig. 1A). Participants with a history of other cognitive impairment and/or psychological disorders were excluded. An MRI safety screen ensured that each participant was safe to enter the bore of the magnet. Participants gave informed consent according to the procedures approved by Boston University Institutional Review Board.

2.2. Assessment

All participants were given a battery of standardized language tests, including the Western Aphasia Battery (WAB-R; Kertesz, 2006) to establish the type and severity of aphasia, the Boston Naming Test (BNT; Goodglass, Kaplan, & Weintraub, 1983) to determine confrontation naming ability, selected subtests of the Psycholinguistic Assessment of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992) to determine specific deficits of access to the semantic system, the Pyramids and Palm Trees (PAPT; Howard & Patterson, 1992) to determine overall soundness of the semantic system, and the Cognitive Linguistic Quick Test (CLQT; Helm-Estabrooks, 2001) to determine the relative contribution of cognitive deficits such as attention and memory to language dysfunction. See Table 2 for full demographic information.

2.3. Treatment

The treatment that was used in this study has previously been shown to be effective in PWA (Kiran et al., 2009) and the treatment results of the patients in the current study are also reported in Sandberg and Kiran (2014). Treatment was carried out in a multiple-baseline

single-subject research design. Each participant was trained on ten abstract words (e.g., *justice*) in a context-category (e.g., *courthouse*), while ten concrete words (e.g., *lawyer*) in the same context-category served as the untrained items used to measure generalization. A second context-category (e.g., *hospital*) served as a control and assignment was counterbalanced across participants. Each participant received therapy twice per week for 2 h each session. Treatment steps included choosing features related to the target, deciding whether the word was abstract or concrete in nature, and generating a synonym for the target. Treatment outcomes were tested using a word generation task in which the patient was asked to generate as many words as possible in 2 min within each category. Patients were encouraged to generate both abstract and concrete words, with definitions and examples provided. Accuracy was based on the percent of predetermined target abstract and concrete words that were generated within each category. Target abstract and concrete words for each category were determined based on responses from a group of healthy individuals in a previous study (Kiran et al., 2009). The treatment and testing protocols were the same as in Kiran et al. (2009) and are reported in more detail in Sandberg and Kiran (2014).

2.4. Functional magnetic resonance imaging

2.4.1. Stimuli and task—The abstract and concrete words used as stimuli in the fMRI task were obtained from the MRC Psycholinguistic Database (Coltheart, 1981) and overlapped with the stimuli used in treatment. Importantly, *t*-tests confirmed that the abstract and concrete stimuli differed on concreteness ($t(55) = 21.64$, $p < .001$) and imageability ($t(57) = 180.29$, $p < .001$), but did not differ on frequency ($t(51) = 1.15$, $p = .26$), familiarity ($t(57) = 1.47$, $p = .15$), letter length ($t(118) = .56$, $p = .58$), or number of syllables ($t(116) = .61$, $p = .54$) (Brown, 1984; Gilhooly & Logie, 1980).

For the fMRI task, we utilized a word judgment (WJ) task. While it was important that the task used in the scanner resemble as close as possible the task used to probe treatment effects (Kiran et al., 2013), using a word generation task was untenable because (a) there is no way to control for the number or quality of generated abstract and concrete words inside the scanner, which introduces the potential for differences in power for specific contrasts in the analysis that may produce results that are unrelated to the underlying mechanism of interest, (b) head motion in the scanner can drive large changes in estimates of functional connectivity, particularly over long distances (Power, Barnes, Snyder, Schlaggar, & Petersen, 2012), thus leading to the potential to overestimate the functional network involved in tasks that include overt speech, and (c) in an intermittent task, such as word generation, modeling the BOLD time course related to the neural/cognitive processes of interest, which is required for task-based functional connectivity in event-related designs (as performed here), could be problematic, especially in PWA with word-finding difficulties. While it might be possible to overcome the technical challenges of estimating functional connectivity in such a task that evokes variable response timing, stimulus-correlated head motion, and an unpredictable number of data points within each condition, we determined that it would be more expeditious to use a modified task that was related to the treatment task, but that could be modeled more effectively with a traditional general linear model. We thus designed the WJ task to directly relate to one of the training steps used during

treatment, which is to determine if the word being trained is abstract or concrete in nature, in order to more closely align the task performed in the fMRI scanner to the semantic training that occurred during treatment. During the WJ task, abstract words (e.g., *justice*) and concrete words (e.g., *lawyer*) and letter strings made up of all consonants (e.g., *rvtsg*) or all vowels (e.g., *aoei*) were randomly presented on a projection screen for 4 s each. Participants determined if each word was abstract or concrete or if each letter string was composed of vowels or consonants and pressed the corresponding button. The words *abstract/concrete* and *vowels/consonants* were shown on the bottom of the screen during experimental and control trials, respectively, to eliminate errors or increased effort related to remembering which option corresponded to which button (see Fig. 1B). Importantly, both the experimental condition and the control condition required visual letter analysis and both required the participant to make a categorical decision and button press. The experimental condition – abstract and concrete words – additionally required word retrieval and semantic processing related to what makes each word either abstract or concrete.

During each fMRI scan (see Fig. 1), participants completed four runs of the WJ task for a total of 60 abstract words, 60 concrete words, and 60 letter strings. The presentation of items within each run was randomized and the presentation of runs was counterbalanced across participants and across sessions. For the first three participants, there were three runs with a total of 50 abstract words, 50 concrete words, and 50 letter strings. E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) was used to present the WJ task in the scanner. Both accuracy and reaction time were recorded for analysis.

2.4.2. Experimental design—An event-related design with pseudorandomized interstimulus intervals (ISI; range 1–3 s) was used, allowing for a fine-grained analysis between conditions (i.e., abstract, concrete) by measuring and analyzing the hemodynamic response for each stimulus. Additionally, three participants served as their own controls by being scanned twice (10 weeks apart) before treatment began (see Fig. 1A). During this control period, we did not force the participants to stop their current therapy, but we confirmed that the treatment was unrelated to our designed intervention. In other words, the stimuli used in the current experiment for both the treatment and the fMRI task were not part of their current therapy, and the target of therapy was different (e.g., syntax, discourse).

2.4.3. Protocol—The fMRI scanning was conducted at the Boston University Center for Biomedical Imaging in a 3 Tesla Philips Achieva MRI scanner with an 8-channel head coil that has been approved for use on humans. Participants receiving treatment immediately were scanned once during pre-testing before treatment sessions began (S1) and again during post-testing after treatment sessions had ended (S2). The participants who served as their own controls first participated in a control scan (S0), then engaged in a no-treatment control period of at least 10 weeks, at the end of which they were scanned (S1) and given the choice to begin treatment. All three control participants chose to participate in treatment and were scanned a third time after treatment ended (S2). Participants first practiced the task outside the scanner on a laptop to decrease performance-related anxiety while in the scanner.

2.4.4. Data collection—High-resolution T1-weighted images were acquired with the following parameters: 140 sagittal slices, 1 mm^3 voxels, 240 × 240 matrix, flip angle = 8°,

fold-over direction = AP, TR = 8.2 ms, TE = 3.8 ms. Blood-oxygen-level-dependent (BOLD) sensitive functional images were collected using the following parameters: 31 axial slices (3 mm thick with 0.3 mm gap), $3mm^3$ voxels, 80×78 matrix, flip angle = 90°, fold-over direction = AP, TR = 2000 ms, TE = 35 ms. The visual stimuli were presented on a screen behind the scanner, which projected to a mirror fitted to the head coil. Padding was used to minimize head motion and corrective optical lenses were used when necessary to correct visual acuity. After bore entry, the magnet was shimmed to achieve maximum homogeneity.

2.4.5. Data analysis—Analyses were focused at the individual level as averaged group analysis can mask significant perilesional activation from individuals with heterogeneous lesions (Meinzer et al., 2012). After data were analyzed at the individual level, individual functional connectivity changes were compared across patients to determine consistent patterns for the group.

2.4.5.1. Preprocessing: Using SPM8 software (Wellcome Trust Centre for Neuroimaging), preprocessing was performed to correct for slice timing differences, correct for movement, remove slow baseline drifts, coregister the structural and functional images, and spatially normalize both structural and functional images to the default MNI template. Slow baseline drifts were filtered out using a high-pass filter with a cutoff of 1/128s, and volumes with large variations in scan-to-scan motion or average global intensity were repaired via linear interpolation using the ArtRepair toolbox for SPM8 (Mazaika, Hoeft, Glover, & Reiss, 2009). Spatial smoothing of the functional data was not performed to minimize the loss of specific activations that can occur due to smoothing (see Meinzer et al., 2012 for a discussion regarding smoothing in patient populations). In addition to these basic steps, lesion maps were drawn by hand using MRIcron software (http://www.cabiatl.com/mricro/) and transformed into binary lesion masks that were used during segmentation to minimize deformities during normalization (Brett, Leff, Rorden, & Ashburner, 2001). SPM8 utilizes unified segmentation, which decreases the lesion effect on normalization; however, lesion masking is still recommended (Andersen, Rapcsak, & Beeson, 2010).

2.4.5.2. General linear model (GLM): After preprocessing, data were entered into a GLM in SPM8 for a first-level analysis at the individual level. Only correct responses were included in the conditions and contrasts of interest, since activation during incorrect responses can reflect maladaptive processing or error monitoring (Postman-Caucheteux et al., 2010). Trials that had incorrect responses were modeled using a separate regressor of no interest. The temporal derivative was included in the model to account for possible stroke-related differences in the hemodynamic response function (HRF) (Friston, Josephs, Rees, & Turner, 1998). Anatomical labels for significant activations were obtained from the Automated Anatomical Labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). Activations were thresholded to control the false discovery rate (FDR) at 0.05. Significant peaks of activation from this first-level GLM were used to define functional regions of interest (fROIs) for the functional connectivity analysis. An extension of this GLM approach was used to analyze pre- versus post-treatment BOLD activations at the individual and group

level, which are not the focus of this study, but will be briefly reviewed in Section 3.3 as a compliment to the functional connectivity results.

2.4.5.3. Functional connectivity

2.4.5.3.1. Functional region of interest definition: FROIs were defined methodically for each individual participant by selecting atlas regions that showed activation at an FDR $p < .05$ level for the [abstract > letter strings] contrast and the [concrete > letter strings] contrast in the GLM during scan S1. This personalized list was then compared against the patient's uncorrected $p < .001$ S1 and S2 activations for the [word (abstract + concrete) > letter strings] contrast. Only regions that contained significant activation at both S1 and S2 were retained as fROIs. Using Marsbar (Brett, Anton, Valabregue, & Poline, 2002), we created fROIs by drawing a 5 mm sphere around the MNI coordinates of the peak voxel within each region for each participant. The same procedures were used to create fROIs for S0 and S1 for the three participants who served as their own controls.

2.4.5.3.2. Connectivity matrix definition: Once fROIs were defined for each participant, the CONN toolbox (Whitfield-Gabrieli & Nieto-Castanon, 2012) was used to perform a task-related (i.e., condition-specific) functional connectivity analysis on the data from each time interval (S0, S1, S2). The following steps were carried out within CONN: First, the GLM from SPM8 (as used for individual activation estimates) was used to define model parameters, including the onsets and durations for each condition. Next, Comp-Cor (Behzadi, Restom, Liau, & Liu, 2007) was used to estimate noise components from non-neuronal sources. The variance in BOLD time series related to the main effects of each condition, head motion, and temporal components extracted from the white matter and CSF was then removed. The residual time series were then weighted by the appropriate HRF-convolved regressor to derive condition-specific time series for functional connectivity analyses. Finally, semipartial correlations were then estimated between each pair of ROIs based on these weighted residuals, resulting in a directed, weighted adjacency (i.e., semipartial correlation) matrix for each experimental condition (abstract, concrete, letter strings) for each time period. Note that CONN automatically converts *r*-values to *z*-values using a Fisher transformation, and these formed the input for subsequent analysis. We used semipartial correlations in order to obtain the unique correlation between a seed region and a target region, controlling for the influence of other regions on the target region.

For each participant, we then created a matrix representing *changes* in correlation strength from pre- to post-treatment ($n = 10$, S2–S1) and during the control period ($n = 3$, S1–S0) for each condition using matrix subtraction. These matrices were treated as graphs (which we refer to hereafter as "difference networks"), with nodes representing brain regions, and edges encoding differences in semipartial correlations between region pairs from pre- to post-treatment. While these "difference networks" do not themselves represent a meaningful physiological network, they instead represent the increases in functional connections in a physiological network, and are amenable to the same types of graph analytic approaches used on simpler functional connectivity networks. The letter-string (control condition) difference network was treated as a baseline and was subtracted from the abstract (trained) difference network and the concrete (generalized) difference network to isolate changes in

functional connectivity above and beyond any "baseline" scan-to-scan changes that may occur.

2.4.5.3.3. Node degree calculation: Node degree, a graph theoretic measure, was calculated for each node (region) in these difference networks to determine regions within each network that have large changes in functional connectivity with other regions from pre- to post-treatment, and which are related to treatment gains both within and across patients.

Node degree is defined as the number of connections or edges (incoming and outgoing), adjacent to a particular node (Sporns, 2011). For each participant, we calculated node degree using Eq. (1), where k is the node degree, i is the node, a_{ij} is the connection from node i to node j, a_{ji} is the connection from node j to node i, and N is the set of all nodes (Rubinov & Sporns, 2010).

$$k = k_i^{out} + k_i^{in} = \sum_{j \in N} a_{ij} + \sum_{j \in N} a_{ji} \quad (1)$$

To calculate node degree we first binarized the previously described difference network, such that *significant* increases in functional connectivity from node i to node j was indicated by $A_{ij} = 1$; functional connections that did not show significant strength increases had matrix entries set to 0. The significance of a change in correlation strength for each connection was determined by calculating 95% confidence intervals for the difference between the two Fisher transformed z-scores (Lane, 2007). For our purposes, only significant increases (i.e., z-scores whose confidence intervals did not overlap) were retained.

In our data set, since nodes correspond to brain regions and connections to changes in (semi-partial) correlation values, a node's degree *in the difference network* provides information about the contribution of an individual brain region to network changes coinciding with treatment outcomes. Specifically, this measure defines *how many* connections between that ROI and the rest of a subject-specific network *increase* in connection strength following treatment. Node degrees were calculated at the individual level and then averaged across participants.

3. Results

3.1. Behavioral results of treatment

Nine of the ten participants improved on the trained abstract words with effect sizes (ES) ranging from 4.60 to 17.53. Of the nine participants who responded to treatment, seven also showed generalization from abstract to concrete words in the same context-category with ES ranging from 1.73 to 7.01 for the untrained concrete words. See Fig. 2 for details. For a more detailed analysis of the behavioral results of treatment, see Sandberg and Kiran (2014). Pre- and post-treatment scores on standardized tests can be found in Supplementary Table 1; however, a detailed analysis and discussion of the relationship between standardized test scores and effect sizes is outside the scope of this paper.

3.2. Behavioral results of fMRI task

All participants performed at above-chance levels on the WJ task in the scanner during each time interval (binomial $p < .001$). Wilcoxon signed rank tests showed that overall accuracy for the group did not change from pre- to post-treatment ($p = .35$). Likewise, overall reaction time for the group did not significantly change after treatment ($p = .20$) (see Supplementary Table 2 for individual data). Because only correct responses were analyzed, all activations are assumed to be related to engagement of the semantic processing network.

3.3. fMRI activation results

As the focus of this paper is the change in functional connectivity after treatment, the BOLD activation results will only briefly be addressed. A more in-depth analysis of activation changes related to this treatment will be presented in a separate paper (Sandberg & Kiran, in preparation). However, it is important to broadly present the BOLD activation results here to provide a reference for the functional connectivity results.

Similar to the connectivity analysis, the pre- vs. post-treatment analysis of BOLD activation was performed at the individual level and then at the group level. In addition to the preprocessing steps mentioned in the methods, volumes were also detrended using a voxel-level linear model of the global signal (LMGS; Macey, Macey, Kumar, & Harper, 2004) so that changes in the global signal that are of no interest would not influence comparisons between pre- and post-treatment sessions. At the individual level, a GLM was implemented as a first-level analysis with contrasts of interest being [post-treatment abstract > pre-treatment abstract] and [post-treatment concrete > pre-treatment concrete]. At the group level, a one-sample t-test was used as a second-level analysis to characterize similarities among responders ($n = 9$) for the [post-treatment abstract > pre-treatment abstract] contrast and among generalizers ($n = 7$) for the [post-treatment concrete > pretreatment concrete] contrast.

All participants exhibited significantly increased activation (FDR $p < 0.05$) from pre- to post-treatment for both abstract and concrete words at the individual level, except P4, who only showed increased activation for abstract words at this level of significance (see Supplementary Fig. 2). As shown in Fig. 3[1] and Table 3, the group level analysis resulted in activation for abstract words in responders in several regions, including but not limited to, those preferential to abstract word processing (e.g., L IFG). For concrete words, group level results showed increased activation in generalizers in several regions, including but not limited to, those preferential to concrete word processing (e.g., L PCN). Overlaps in increased activation between abstract and concrete words included L SMG, L IPL, L PCN, and R SFG. However, these patterns of increased activation were only significant at the uncorrected $p < 0.001$ level.

---

[1]Fig. 3 and Supplementary Fig. 2 were created using Caret software http://brainmap.wustl.edu.

3.4. Functional connectivity results

The results here focus on specific patterns of increases in functional connectivity across participants, organized around our main hypothesis: (a) control period (i.e., S1–S0) for the three control participants, (b) direct training effects, and (c) generalization effects.

3.4.1. Control period (n = 3)—In order to probe whether functional connectivity changes in the absence of targeted therapy, a subset of three participants (P8, P9, and P10) served as their own controls, participating in a 10 week control period before beginning therapy. It is important to note here that all three participants exhibited stable baseline performance on the word generation task during the control period (from S0 to S1), but showed significant increases in generative naming performance after treatment (from S1 to S2), improving on both the trained abstract words and generalizing to the untrained concrete words.

At the individual level, pre-treatment (semi-partial) correlation matrices were subtracted from post-treatment correlation matrices for each condition (abstract, concrete). Increases in correlation for each condition-specific connection for each participant that pass the 95% confidence interval threshold are visualized as heat maps in Supplementary Fig. 3. Unexpectedly, all three control participants showed increased functional connectivity after the control period. The implications for this finding will be addressed in the Discussion. Interestingly, the patterns of change for each participant are very different, and notably, each control participant exhibits a different pattern of change after treatment than after the control period (see Supplementary Fig. 4). Importantly, consistencies across participants become apparent with the analysis of node degree in the appropriate difference networks.

Node degrees in the difference networks were calculated at the individual level (see Supplementary Fig. 5 for individual graphs) and then averaged across participants. Only regions that showed significant changes in at least two of the three participants were entered into the analysis. Fig. 4 visually depicts the node degree for both the abstract and concrete difference networks.[2] For the control period, the region with the highest node degree – representing an increase in functional connectivity – for the abstract word network was L SFG, and for the concrete word network was L MFG, brain areas which are involved in attention in addition to semantic processing (Binder et al., 2009; Fan, McCandliss, Fossella, Flombaum, & Posner, 2005; Price, 2012).

Because the control participants also received treatment, we can directly compare the control period versus the treatment period for these participants. Importantly, there are regions that show increases in connectivity across participants regardless of whether or not treatment was given, including L SupMed and L MFG. Interestingly, these regions also tend to have higher node degrees than other regions in the difference networks. Regions that show common increases in functional connectivity between the abstract and concrete networks after treatment, but not the control period include bilateral AG and L IFGtri. Regions specific to the control period for both abstract and concrete networks include L SFG, L preC, L SMA, and L pMTG, although L pMTG also showed direct training effects.

---

[2]Node degree brain maps in Figs. 4 and 5 were generated with the BrainNet Viewer (Xia, Wang, & He, 2013, http://www.nitrc.org/projects/bnv/).

3.4.2. Direct training (n = 9)—Increases in correlation for condition-specific connections for each participant that pass the 95% confidence interval threshold are visualized as heat maps in Supplementary Fig. 5. At the individual level, all participants showed increases in functional connectivity from pre- to post-treatment for the abstract difference network, which we will refer to as the *trained abstract difference network*. Again, each participant shows a different pattern of change after treatment, and consistencies across participants become apparent with the analysis of node degree.

The top panel of Fig. 5 visually depicts node degree for the trained abstract difference network with graphs highlighting the regions for which at least six of the nine participants who responded to treatment showed significant increases in functional connection strength. Regions with high node degree but which are only seen in a few patients (e.g., R SMA had a node degree of 8, but was only seen in one participant) have limited value when attempting to determine consistent, meaningful treatment-related changes in relevant brain networks. Thus, regions for which a substantial fraction – in the present study at least two-thirds – of the responders showed increases are thought to be more informative. Using this criterion, the region with the highest node degree in the trained abstract difference network is the left inferior frontal gyrus *pars triangularis* (L IFGtri).

3.4.3. Generalization (n = 7)—Increases in correlation for condition-specific connections for each participant that pass the 95% confidence interval threshold are visualized in Supplementary Fig. 5. At the individual level, all participants showed significant changes in functional connectivity from pre- to post-treatment for the concrete word difference network, which we will refer to as the *generalized concrete difference network* for participants who generalized and as the *non-generalized concrete difference network* for those who did not generalize. As with direct training effects, each participant shows a different pattern of functional connectivity changes after treatment, and the analysis of node degree better characterizes consistencies across participants, and additionally identifies differences between participants who generalized versus those who did not.

The middle panel of Fig. 5 depicts node degree for the generalized concrete difference network, with plots highlighting the regions for which at least five of the seven participants who generalized in treatment showed significant increases in functional connectivity. Using our two-thirds criterion, the regions with the highest node degree in the generalized concrete difference network were L SupMed and R IFGtri.

3.4.4. Non-generalizers (n = 2)—For comparison, we also examined increases in functional connectivity in the non-generalized concrete difference network. Since there were only two participants who did not generalize, only regions for which both participants showed increases were considered. The regions with the highest node degree in the nongeneralized concrete difference network were L MFG and R IFGorb. See the bottom panel of Fig. 5 for details.

3.4.5. Lateralization of functional connectivity changes—As an additional post hoc analysis we calculated two indices – a Bilaterality Index (BI) and a Laterality Index (LI) – to determine if connection changes were lateralized, and if so, which hemisphere was playing a

larger role in treatment-induced neuroplasticity for both direct training and generalization effects of treatment. The BI and LI were calculated for each participant and then averaged across participants to determine similarities and differences among the trained abstract difference network, the generalized concrete difference network and the non-generalized concrete difference network to better characterize differences between generalizers and non-generalizers.

The BI compares the ratio of cross-hemisphere connections that are increasing in correlation strength to the ratio of within-hemisphere connections that are increasing in correlation strength (see Eq. (2)). For each participant, the number of left-to-left (LL), right-to-right (RR), left-to-right (LR) and right-to-left (RL) hemisphere connections that were significantly increasing in correlation strength were summed and then entered into Eq. (2). A positive BI value indicates more cross-hemisphere connection changes and a negative BI value indicates more within-hemisphere connection changes.

$$BI = \frac{(LR+RL) - (LL+RR)}{(LR+RL+LL+RR)} \quad (2)$$

The LI compares the ratio of left-hemisphere (LH) connections that are increasing in correlation strength to the ratio of right-hemisphere (RH) connections that are increasing in correlation strength (see Eq. (3)). A positive LI value indicates a LH bias and a negative LI value indicates a RH bias. A bias of ±0.2 has been accepted as a reasonable threshold for laterality (Seghier, 2008) and will be used as a threshold for both BI and LI in interpreting our results. A paired t-test confirmed no significant difference between the number of LH versus RH fROIs across participants ($t = 1.97$, $p = .09$). Thus, we can assume that a laterality bias is not simply the result of the number of LH versus RH fROIs included in the functional connectivity matrices.

$$LI = \frac{(LL - RR)}{(LL+RR)} \quad (3)$$

Both generalizers and non-generalizers showed an average negative BI beyond the −0.2 threshold for increased functional connectivity in the trained abstract network, indicating a bias for within-hemisphere connection changes related to direct training effects. However, only generalizers showed an average negative BI beyond the −0.2 threshold for increased functional connectivity in the generalized concrete difference network, similar to the trained abstract network, indicating a bias for within-hemisphere connection changes related to generalization. For non-generalizers, the average BI value for the concrete network did not pass the −0.2 threshold, indicating a difference between the abstract and concrete networks related to the difference between direct training and generalization effects of treatment for this group (see Fig. 6A).

Upon closer examination of within-hemisphere connections, generalizers exhibited an average positive LI beyond the 0.2 threshold for both the trained abstract network and the generalized concrete difference network, indicating a LH bias for increased functional

connectivity related to both direct training and generalization effects. On the other hand, non-generalizers exhibited an average positive LI beyond the 0.2 threshold for the trained abstract network, but an average negative LI beyond the threshold for the non-generalized concrete difference network, indicating a LH bias for increased functional connectivity related to direct training, but a RH bias for increased functional connectivity related to a failure to generalize in treatment (see Fig. 6B). While there is individual variability (see Supplementary Fig. 6), the general trend remains consistent.

4. Discussion

The aim of this study was to examine changes in functional connectivity patterns associated with both training and generalization effects of a word finding treatment in persons with aphasia. We found that increases in functional connectivity were consistent with BOLD activation changes and with the effects of treatment, including (a) differences between the control period and the treatment period for the participants who served as their own controls, (b) similarities between direct training and generalization effects of treatment, and (c) differences between generalization and non-generalization effects of treatment.

4.1. Control period

We hypothesized that there would be no significant increases in functional connectivity after the control period. Unexpectedly, all three participants who served as their own controls showed changes in functional connectivity after the control period. For both abstract and concrete difference networks, L SFG, L preC, and L SMA – regions involved in attention and subvocal rehearsal (Fan et al., 2005; Price, 2012) – were unique to control period changes. It is possible that changes that occur during the control period are related to practice effects (i.e., differences in functional connectivity due to familiarity with the task); however, we do not believe this is the case, since all participants practice the task on a laptop until they are comfortable with the task prior to scanning.

It is also possible that the changes noted during the control period were related to the unrelated therapies these participants were receiving. Although they were not receiving word-finding treatment and were not exposed to the treatment stimuli, sentence comprehension and discourse therapy necessarily involve activation of the semantic system. There were overlaps between the treatment and control periods in regions that showed increased connectivity, including L SupMed and L MFG. While these regions are considered part of the concrete network, parts of SupMed and MFG overlap with the default mode network (Binder et al., 2009; Buckner, Andrews-Hanna, & Schacter, 2008), and parts of MFG are also thought to be involved with working memory (Curtis & D'Esposito, 2003). It would be interesting to examine whether the changes noted would be any different than therapy in other cognitive domains or exposure to everyday conversation.

Importantly, the regions that showed increased connectivity consistently across participants after the control period were qualitatively different than those related to the treatment period. For both abstract and concrete difference networks, connectivity changes in L IFG and bilateral AG – regions important for semantic processing (Binder et al., 2009; Price, 2012) – were unique to the treatment period. These results suggest that, for these

participants, successful treatment may coincide with functional connectivity changes in regions within the semantic network, and that these same changes may not arise simply as a function of time following the onset of the stroke. Additional work will be necessary to determine if such changes are truly specific to successful treatment (i.e., involving additional patients scanned pre- and post-control period), and if such changes are causally related to successful outcomes.

4.2. Direct training effects

We hypothesized that direct training effects of treatment would coincide with increased functional connectivity in the prescribed abstract word network, specifically L IFG, L STG, and/or L MTG. Indeed, across participants who improved on the trained abstract words (n = 9), increased functional connectivity was observed in the prescribed abstract network and in additional regions, with some overlap among networks (see Section 4.4). Importantly, the highest node degree for the trained abstract difference network was in L IFGtri, indicating that this may be an important node for plasticity within in the prescribed abstract network. Similarly, this region showed increased BOLD signal in the group analysis, suggesting that L IFGtri is not only gaining more connections, but is increasing activation and perhaps influence over established and newly connected regions, further supporting the role of this region in the improvement of abstract word processing,

4.3. Generalization effects

We hypothesized that generalization effects of treatment would coincide with increased functional connectivity in the prescribed concrete word network, specifically bilateral AG, L FFG, L PHG, L SupMed, L SFG, L MFG, L PCN, and/or L PCC. Across participants who generalized to the untrained concrete words in the same context-category (n = 7), increased functional connectivity was observed in the concrete difference network, with the highest node degree in L SupMed and R IFGtri. L SupMed is one of the regions in the prescribed concrete network, but R IFGtri is not. However, R IFGtri has been indicated as a supporting node to L IFGtri in aphasia recovery (Turkeltaub, Messing, Norise, & Hamilton, 2011), and our results remain consistent with a possible contribution of this region in post-treatment networks. While L PCN, another region in the prescribed concrete network, did not have the highest node degree in the concrete difference network for this group, it did show connectivity changes for 5 of the 7 generalizers. Similarly, L PCN showed increased BOLD activation for generalizers as a group. This suggests that while L PCN is not increasing its connectivity with a high number of regions, it is increasing its activation, possibly influencing established connections more. These results partially support our hypothesis and underline the complimentary nature of BOLD activation and functional connectivity analyses.

Unexpectedly, the two participants who did not generalize in treatment also showed increased functional connectivity in the prescribed concrete network, albeit in different nodes than generalizers. For non-generalizers, the highest node degree in the concrete difference network was in L MFG and R IFGorb. In order to tease apart the differences between generalizers and nongeneralizers, the remainder of this discussion will focus on the comparison between treatment outcomes.

4.4. Differences in treatment outcomes

Because our participants did not all respond to treatment in the same way, we were able to compare direct training effects, generalization effects, and non-generalization effects. When considering regions of increased functional connectivity with high agreement among participants, we see considerable overlap between the trained abstract difference network and the generalized concrete difference network, including L AG, R AG, L IFGtri, R IFGtri, L IFGorb, R IFGorb, L MFG, L SupMed, L PCN, and L pMTG. Of these regions, only L IFGtri and bilateral AG are not also found in networks other than the trained abstract difference network and the generalized concrete difference network (see Fig. 6C). Thus, the regions that are most important for success in treatment appear to be L IFGtri, with higher node degree for the trained abstract difference network than the generalized concrete difference network, and bilateral AG, with higher node degree for the generalized concrete difference network than the trained abstract difference network. This is in line with previous neuroimaging studies showing preferential activation for abstract words in L IFG and preferential activation for concrete words in bilateral AG (Binder et al., 2009; Wang et al., 2010).

Interestingly, we also see regions of increased functional connectivity that overlap among all treatment conditions, including L MFG, L PCN, R IFGtri, and R IFGorb. Importantly, L MFG also overlaps with the increases in functional connectivity seen after the control period. Thus, L MFG may simply indicate baseline scan-to-scan changes, while L PCN, R IFGtri, and R IFGorb may be related to a general effect of treatment, but not necessarily success in treatment. It is important to note here that practice effects were assumed to be mitigated by practice with the task outside the scanner prior to the first scan. These unexpected changes present an interesting opportunity for future work examining scan-to-scan fluctuations in functional connectivity that occur over time in both healthy older adults and PWA. Recent neuroimaging work has shown good test–retest reliability on graph theoretical measures for resting state (Wang et al., 2011) and task-based (e.g., working memory, emotion processing) functional connectivity (Cao et al., 2014) in healthy young adults. Guo et al. (2012) also found good test–retest reliability on graph theoretical measures for resting state functional connectivity in healthy older adults. However, test–retest reliability for task-based functional connectivity specific to language tasks in healthy older adults, let alone PWA, has not yet been established.

Perhaps more compelling are the bilaterality and laterality biases related to treatment outcome, such that success in treatment coincides with not only more within-hemisphere changes, but specifically left-lateralized changes in connectivity, while failure to generalize coincides with right-lateralized changes in connectivity. While there was, naturally, individual variability, this lateralization effect does not appear to be related to the number of LH versus RH fROIs (see Section 3.4.5), nor to lesion size, since LI was not correlated with lesion size for either the abstract ($r = -.35$, $p = .36$) or concrete ($r = -.50$, $p = .17$) networks.

In addition to these results, there are several aspects of this study that are innovative and represent advancements in the ability to account for individual variability while performing a group-level analysis. First, we used fROIs based on condition-specific activations for each

individual. This allows not only the use of many ROIs, but also a varying number of ROIs for each individual, which is uncommon in functional connectivity analyses. We also used a graph theoretical approach to functional connectivity analysis which allowed a comparison across participants while taking into account individual variability. Node degree has not yet been explored in the examination of changes in functional connectivity associated with treatment gains, but is particularly well-suited to explore changes in functional connectivity at both the individual level and across participants, and collapses results to the familiar level of individual regions of interest. Further, we applied a laterality index (and devised and applied a similar index of bilaterality) to a functional connectivity analysis, which proved to be effective in teasing apart differences between generalizers and non-generalizers.

4.5. Limitations

One limitation of this study was the small number of participants who served as their own controls. While their data was informative, the comparison of changes that occur after a control period versus after treatment would be more powerful with additional participants. Similarly, a healthy control group would have provided information regarding normal scan-to-scan fluctuations that occur over time, providing a context for the current results. Additionally, we only had two participants who did not generalize in treatment. It would be more informative to be able to compare a more balanced group of generalizers versus non-generalizers; however, since this treatment was designed to promote generalization, it would be difficult to amass a large group of non-generalizers.

Another limitation of this study is that the Word Judgment task used in the scanner was different than the word generation task used to probe direct training and generalization effects during treatment; however, it is important to note that the task used in the scanner was a training step used in treatment. While it is preferable to exactly mirror the task used in the scanner to the task used to probe treatment effects in order to create a stronger link between the behavioral and neurophysiological outcomes of treatment, we believe the WJ task allowed us to capture changes in functional connectivity that reflect a fundamental change in the semantic network that resulted from treatment, not just increased generative naming. Specifically, because performance on the WJ task remained relatively stable from pre- to post-treatment (and assuming no change in the noise structure of the residual time series from pre- to post-treatment), we should have similar power before and after treatment for our comparisons, meaning that whatever change we are measuring is not reflecting an increase in data points, but an actual change in the dynamics of semantic processing. Additionally, by grouping and analyzing the data based upon treatment outcomes, we have imposed a link between the behavioral and neurophysiological outcomes. We find it especially convincing that although the task used in the scanner did not exactly mirror the task used to probe treatment effects, functional connectivity changes did mirror treatment outcomes.

5. Conclusion

In this study, we aimed to shed light on the underlying neural mechanism of generalization from abstract to concrete words by comparing functional connectivity patterns of neural

systems engaged in abstract and concrete word processing in PWA before and after a word generation treatment. Like Abutalebi et al. (2009), Sarasso et al. (2010), and Vitali et al. (2010), we found that direct training effects coincided with increased functional connectivity. Specifically, we found that across PWA who responded to treatment, L IFGtri appears to be an important node of plasticity, increasing its activation and communication with other brain areas in response to treatment when abstract words are trained. Like Vitali et al. (2010), we observed similarities between the trained and generalized difference networks, with L IFG and bilateral AG surfacing as important nodes of plasticity for success in treatment. Additionally, we observed differences between the generalized and non-generalized difference networks, both in terms of regional node degree and hemispheric laterality, indicating a distinct effect of success in treatment on connectivity changes. Because training abstract words is not the only way to promote generalization, future research testing this notion in other modes of generalization will help uncover the neural mechanisms that promote effective relearning in persons with aphasia. This information is expected to help guide the development of more effective and efficient treatments for aphasia.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Abutalebi J, Rosa PA, Tettamanti M, Green DW, Cappa SF. Bilingual aphasia and language control: A follow-up fMRI and intrinsic connectivity study. Brain and Language. 2009; 109(2–3):141–156. http://dx.doi.org/10.1016/j.bandl.2009.03.003, pii: S0093-934X(09)00040-6. [PubMed: 19427522]

Andersen SM, Rapcsak SZ, Beeson PM. Cost function masking during normalization of brains with focal lesions: Still a necessity? Neuroimage. 2010; 53(1):78–84. http://dx.doi.org/10.1016/j.neuroimage.2010.06.003. [PubMed: 20542122]

Behzadi Y, Restom K, Liau J, Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage. 2007; 37(1):90–101. http://dx.doi.org/10.1016/j.neuroimage.2007.04.042. [PubMed: 17560126]

Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebral Cortex. 2009 http://dx.doi.org/10.1093/cercor/bhp055.

Brett, M.; Anton, JL.; Valabregue, R.; Poline, JB. Paper presented at the human brain mapping. Japan: Sendai; 2002. Region of interest analysis using an SPM toolbox.

Brett M, Leff AP, Rorden C, Ashburner J. Spatial normalization of brain images with focal lesions using cost function masking. Neuroimage. 2001; 14:486–500. [PubMed: 11467921]

Brown GA. A frequency count of 190,000 words in the London-Lund Corpus of English Conversation. Behavior Research Methods, Instruments, & Computers. 1984; 16(6):502–532. http://dx.doi.org/10.3758/BF03200836.

Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network. Annals of the New York Academy of Sciences. 2008; 1124(1):1–38. http://dx.doi.org/10.1196/annals.1440.011. [PubMed: 18400922]

Cao H, Plichta MM, Schäfer A, Haddad L, Grimm O, Schneider M, Tost H. Test–retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing, and resting state. Neuroimage. 2014; 84:888–900. http://dx.doi.org/10.1016/j.neuroimage.2013.09.013. [PubMed: 24055506]

Coltheart M. The MRC psycholinguistic database. Quarterly Journal of Experimental Psychology. 1981; 33A:497–505.

Cornelissen K, Laine M, Tarkiainen A, Jarvensivu T, Martin N, Salmelin R. Adult brain plasticity elicited by anomia treatment. Journal of Cognitive Neuroscience. 2003; 15(3):444–461. http://dx.doi.org/10.1162/089892903321593153. [PubMed: 12729495]

Cramer SC. Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. Annals of Neurology. 2008; 63(3):272–287. [PubMed: 18383072]

Curtis CE, D'Esposito M. Persistent activity in the prefrontal cortex during working memory. Trends in Cognitive Sciences. 2003; 7(9):415–423. http://dx.doi.org/10.1016/S1364-6613(03)00197-9. [PubMed: 12963473]

Fan J, McCandliss BD, Fossella J, Flombaum JI, Posner MI. The activation of attentional networks. Neuroimage. 2005; 26(2):471–479. http://dx.doi.org/10.1016/j.neuroimage.2005.02.004. [PubMed: 15907304]

Fridriksson J, Richardson JD, Fillmore P, Cai B. Left hemisphere plasticity and aphasia recovery. Neuroimage. 2012; 60(2):854–863. http://dx.doi.org/10.1016/j.neuroimage.2011.12.057. [PubMed: 22227052]

Friston KJ, Josephs O, Rees G, Turner R. Nonlinear event-related responses in fMRI. Magnetic Resonance in Medicine. 1998; 39(1):41–52. http://dx.doi.org/10.1002/mrm.1910390109. [PubMed: 9438436]

Gilhooly KJ, Logie RH. Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. Behavior Research Methods and Instrumentation. 1980; 12:395–427.

Goodglass, H.; Kaplan, E.; Weintraub, S. Boston naming test. Philadelphia: Lea & Febiger; 1983.

Guo C, Kurth F, Zhou J, Mayer E, Eickhoff S, Kramer J, Seeley W. One-year test–retest reliability of intrinsic connectivity network fMRI in older adults. Neuroimage. 2012; 61(4):1471–1483. http://dx.doi.org/10.1016/j.neuroimage.2012.03.027. [PubMed: 22446491]

Helm-Estabrooks, N. Cognitive Linguistic Quick Test. San Antonio: The Psychological Corporation; 2001.

Howard, D.; Patterson, K. The Pyramids and Palm Trees Test. Bury St. Edmunds: Thames Valley Test Company; 1992.

Kay, J.; Lesser, RP.; Coltheart, M. The Psycholinguistic Assessment of Language Processing in Aphasia (PALPA). Hove, UK: Erlbaum; 1992.

Kertesz, A. Western Aphasia Battery – Revised (WAB-R). Harcourt Assessment, Inc.; 2006.

Kiran S, Ansaldo A, Bastiaanse R, Cherney LR, Howard D, Faroqi-Shah Y, Thompson CK. Neuroimaging in aphasia treatment research: Standards for establishing the effects of treatment. Neuroimage. 2013; 76:428–435. http://dx.doi.org/10.1016/j.neuroimage.2012.10.011. [PubMed: 23063559]

Kiran S, Sandberg C, Abbott K. Treatment for lexical retrieval using abstract and concrete words in persons with aphasia: Effect of complexity. Aphasiology. 2009; 23:835–853. [PubMed: 19816590]

Lane DM. HyperStat online statistics textbook. 2007

Macey PM, Macey KE, Kumar R, Harper RM. A method for removal of global effects from fMRI time series. Neuroimage. 2004; 22(1):360–366. http://dx.doi.org/10.1016/j.neuroimage.2003.12.042. [PubMed: 15110027]

Marcotte K, Perlbarg V, Marrelec G, Benali H, Ansaldo AI. Default-mode network functional connectivity in aphasia: Therapy-induced neuroplasticity. Brain and Language. 2013; 124(1):45–55. http://dx.doi.org/10.1016/j.bandl.2012.11.004. [PubMed: 23274798]

Mazaika, P.; Hoeft, F.; Glover, GH.; Reiss, AL. Methods and software for fMRI analysis for clinical subjects; Paper presented at the human brain mapping; 2009.

Meinzer M, Beeson PM, Cappa S, Crinion J, Kiran S, Saur D, Thompson CK. Neuroimaging in aphasia treatment research: Consensus and practical guidelines for data analysis. Neuroimage. 2012 http://dx.doi.org/10.1016/j.neuroimage.2012.02.058, pii: S1053-8119(12)00236-4.

Meinzer M, Flaisch T, Breitenstein C, Wienbruch C, Elbert T, Rockstroh B. Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia. Neuroimage. 2008; 39(4):2038–2046. http://dx.doi.org/10.1016/j.neuroimage.2007.10.008, pii: S1053-8119(07)00930-5. [PubMed: 18096407]

Meinzer M, Flaisch T, Obleser J, Assadollahi R, Djundja D, Barthel G, Rockstroh B. Brain regions essential for improved lexical access in an aged aphasic patient: a case report. BMC Neurology. 2006; 6:28. http://dx.doi.org/10.1186/1471-2377-6-28. [PubMed: 16916464]

Postman-Caucheteux WA, Birn RM, Pursley RH, Butman JA, Solomon JM, Picchioni D, Braun AR. Single-trial fMRI shows contralesional activity linked to overt naming errors in chronic aphasic patients. Journal of Cognitive Neuroscience. 2010; 22(6):1299–1318. [PubMed: 19413476]

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage. 2012; 59(3):2142–2154. http://dx.doi.org/10.1016/j.neuroimage.2011.10.018. [PubMed: 22019881]

Price CJ. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage. 2012; 62(2):816–847. http://dx.doi.org/10.1016/j.neuroimage.2012.04.062. [PubMed: 22584224]

Rochon E, Leonard C, Burianova H, Laird L, Soros P, Graham S, Grady C. Neural changes after phonological treatment for anomia: An fMRI study. Brain and Language. 2010; 114(3):164–179. http://dx.doi.org/10.1016/j.bandl.2010.05.005. [PubMed: 20547416]

Rubinov M, Sporns O. Complex network measures of brain connectivity: Uses and interpretations. Neuroimage. 2010; 52(3):1059–1069. http://dx.doi.org/10.1016/j.neuroimage.2009.10.003. [PubMed: 19819337]

Sami S, Miall R. Graph network analysis of immediate motor-learning induced changes in resting state BOLD. Frontiers in Human Neuroscience. 2013; 7:166. http://dx.doi.org/10.3389/fnhum.2013.00166. [PubMed: 23720616]

Sandberg C, Kiran S. How justice can affect jury: Training abstract words promotes generalisation to concrete words in patients with aphasia. Neuropsychological Rehabilitation. 2014:1–32. http://dx.doi.org/10.1080/09602011.2014.899504.

Sandberg CW, Kiran S. Changes in BOLD signal after abstract word generation treatment in aphasia. (in preparation).

Sarasso S, Santhanam P, Määtta S, Poryazova R, Ferrarelli F, Tononi G, Small SL. Non-fluent aphasia and neural reorganization after speech therapy: Insights from human sleep electrophysiology and functional magnetic resonance imaging. Archives Italiennes de Biologie. 2010; 148(3):271–278. [PubMed: 21175013]

Schneider, W.; Eschman, A.; Zuccolotto, A. E-Prime user's guide. Pittsburgh: Psychology Software Tools Inc.; 2002.

Seghier ML. Laterality index in functional MRI: Methodological issues. Magnetic Resonance Imaging. 2008; 26(5):594–601. http://dx.doi.org/10.1016/j.mri.2007.10.010. [PubMed: 18158224]

Sporns, O. Networks of the brain. Cambridge, Massachusetts: MIT Press; 2011.

Thompson C, Shapiro L, Kiran S, Sobecks J. The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The complexity account of treatment efficacy (CATE). Journal of Speech Language and Hearing Research. 2003; 46(3):591–607.

Turkeltaub PE, Messing S, Norise C, Hamilton RH. Are networks for residual language function and recovery consistent across aphasic patients? Neurology. 2011; 76(20):1726–1734. http://dx.doi.org/10.1212/WNL.0b013e31821a44c1, pii: 76/20/1726. [PubMed: 21576689]

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage. 2002; 15(1):273–289. http://dx.doi.org/10.1006/nimg.2001.0978. [PubMed: 11771995]

van Hees S, McMahon K, Angwin A, de Zubicaray G, Read S, Copland DA. A functional MRI study of the relationship between naming treatment outcomes and resting state functional connectivity in post-stroke aphasia. Human Brain Mapping. 2014; 35(8):3919–3931. http://dx.doi.org/10.1002/hbm.22448. [PubMed: 24453137]

Vitali P, Tettamanti M, Abutalebi J, Ansaldo AI, Perani D, Cappa SF, Joanette Y. Generalization of the effects of phonological training for anomia using structural equation modelling: A multiple single-case study. Neurocase. 2010; 16(2):93–105. http://dx.doi.org/10.1080/13554790903329117, pii: 917446096. [PubMed: 19967599]

Wang J, Conder JA, Blitzer DN, Shinkareva SV. Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. Human Brain Mapping. 2010; 31(10):1459–1468. [PubMed: 20108224]

Wang J-H, Zuo X-N, Gohel S, Milham MP, Biswal BB, He Y. Graph theoretical analysis of functional brain networks: Test–retest evaluation on short- and long-term resting-state functional MRI data. PloS One. 2011; 6(7):e21976. http://dx.doi.org/10.1371/journal.pone.0021976. [PubMed: 21818285]

Whitfield-Gabrieli S, Nieto-Castanon A. Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. Brain Connect. 2012; 2(3):125–141. http://dx.doi.org/10.1089/brain.2012.0073. [PubMed: 22642651]

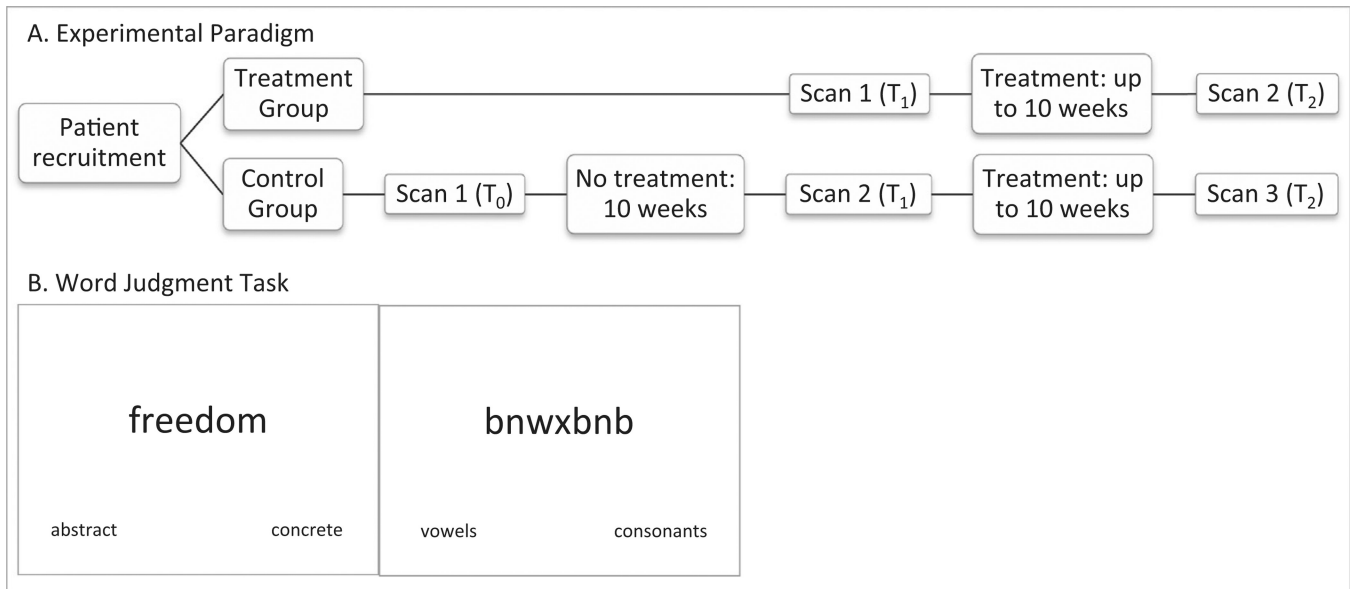Xia M, Wang J, He Y. BrainNet Viewer: A network visualization tool for human brain connectomics. PloS One. 2013; 8(7):e68910. http://dx.doi.org/10.1371/journal.pone.0068910. [PubMed: 23861951]
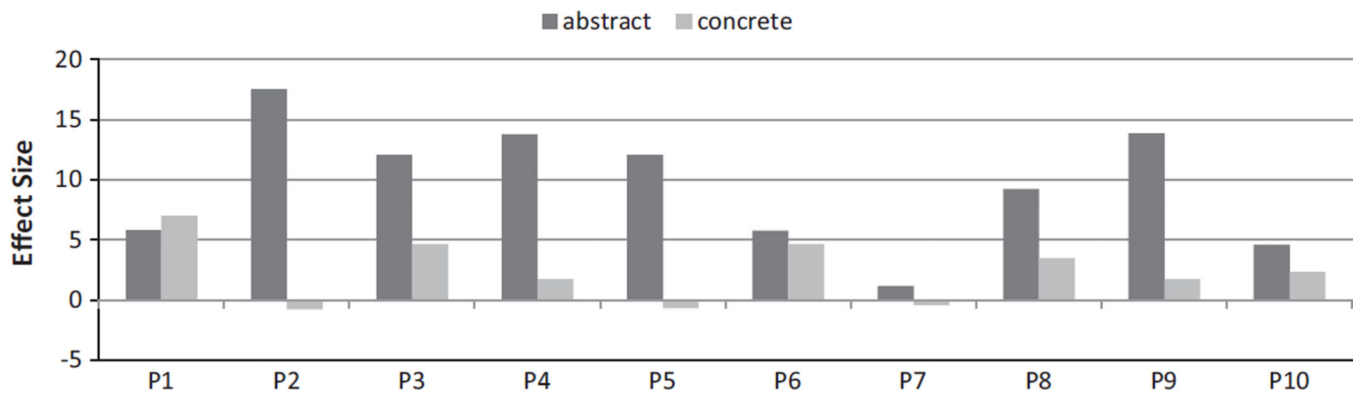
Fig. 1.
Experiment details.

Fig. 2.
Treatment results. The abstract effect size is related to direct training, the concrete effect size is related to generalization. Note that P7 did not respond to treatment (ES < 4) while P2 and P5 improved on abstract words, but did not generalize to concrete words (negative ES).
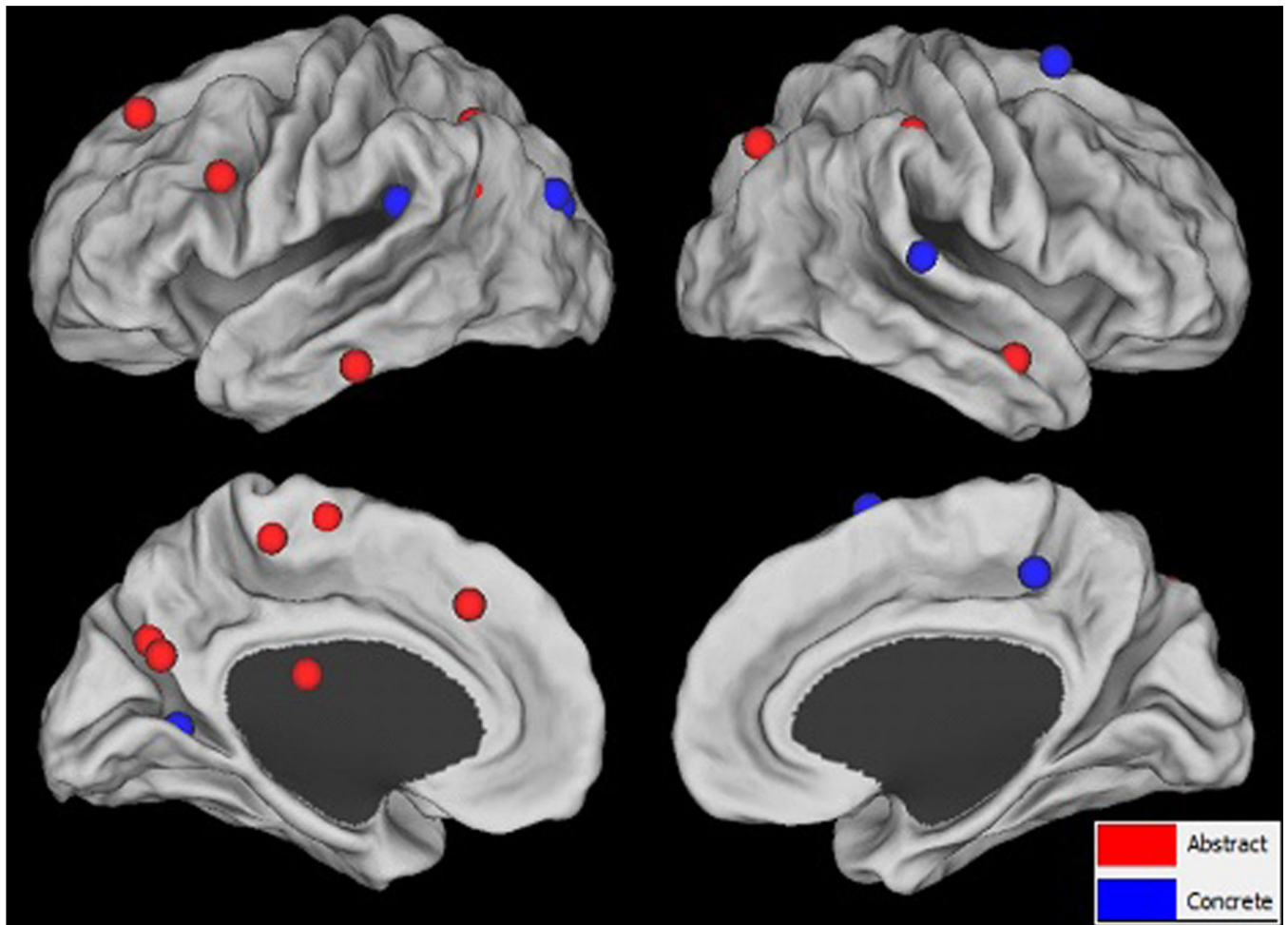
Fig. 3.
Changes in BOLD signal from pre- to post-treatment for abstract and concrete words at the group level. Red spheres indicate peaks of activation for the one-sample *t*-test of the [post-treatment abstract > pre-treatment abstract] contrast for the group of responders (*n* = 9). Blue spheres indicate peaks of activation for the one-sample *t*-test of the [post-treatment concrete > pre-treatment concrete] contrast for the group of generalizers (*n* = 7). All results shown are significant at the uncorrected *p* < 0.001 level. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
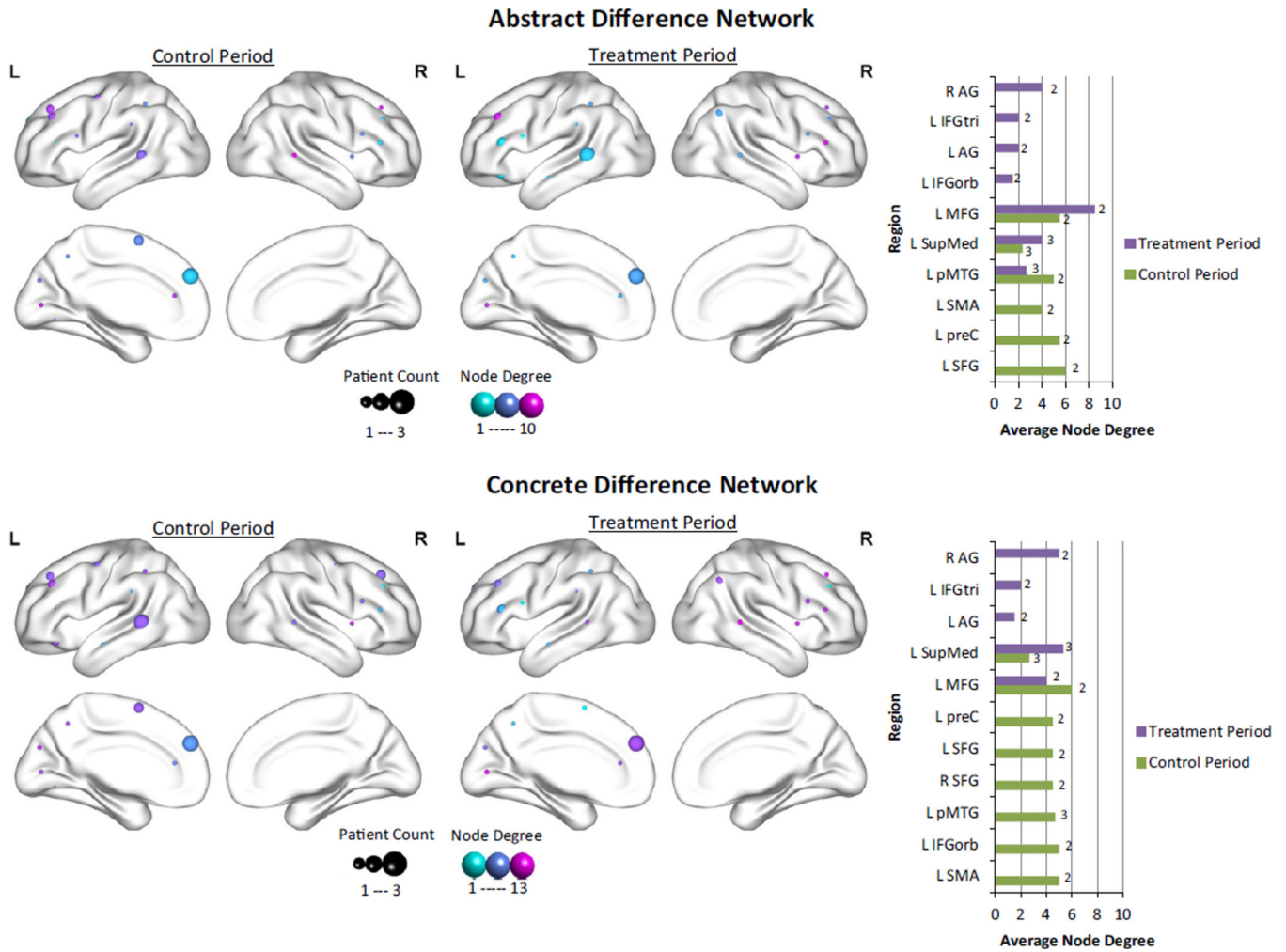
Fig. 4.
Comparison of node degree of fROIs between the control and treatment periods in the abstract and concrete change networks. This figure illustrates the average node degree for increases in connectivity after the control and treatment periods for fROIs in the abstract difference network (top panel) and the concrete difference network (bottom panel) for the three participants who served as their own controls. The size of each sphere represents the number of participants who show significant increases in connectivity for that region, while the color of the sphere represents the average node degree. Higher values are more purple, lower values are more turquoise. The bar graphs highlight the differences between the control (green) and treatment (purple) periods for the majority (at least 2/3) of control participants. The number of participants who showed increased connectivity for each region is provided to the right of each bar. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Trained Abstract Network

Generalized Concrete Network

Non-Generalized Concrete Network
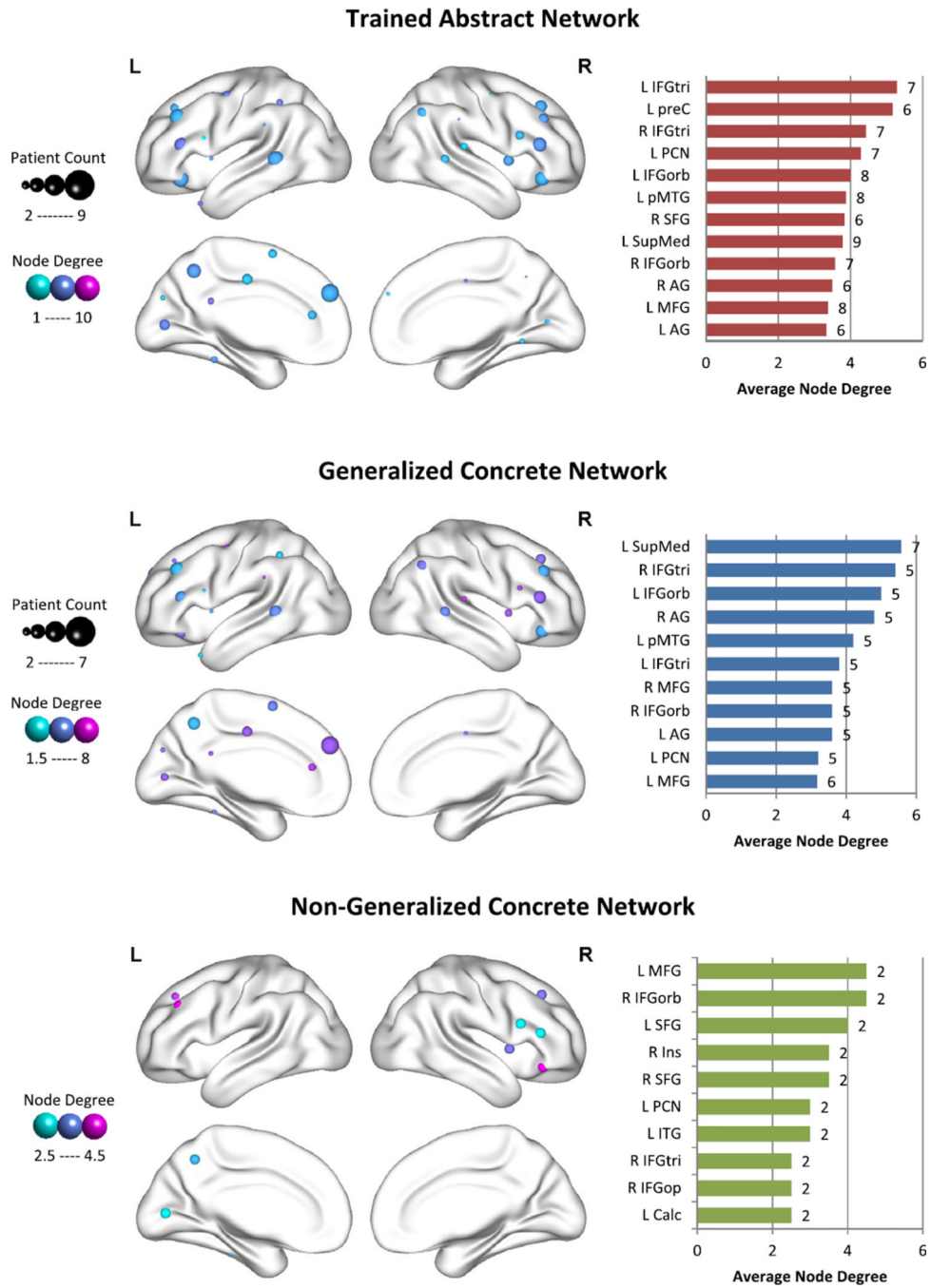
Fig. 5.
Node degree of fROIs in the trained abstract network, generalized concrete network, and non-generalized concrete network. This figure illustrates the average node degree for increases in connectivity for fROIs in the trained abstract network (top panel), generalized concrete network (middle panel), and non-generalized concrete network (bottom panel). The size of each sphere represents the number of participants who show significant increases in connectivity for that region, while the color of the sphere represents the average node degree. Higher values are more purple, lower values are more turquoise. The bar graphs

highlight the regions with the highest node degree for the majority (at least 2/3) of participants. The number of participants who showed increased connectivity for each region is provided to the right of each bar. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
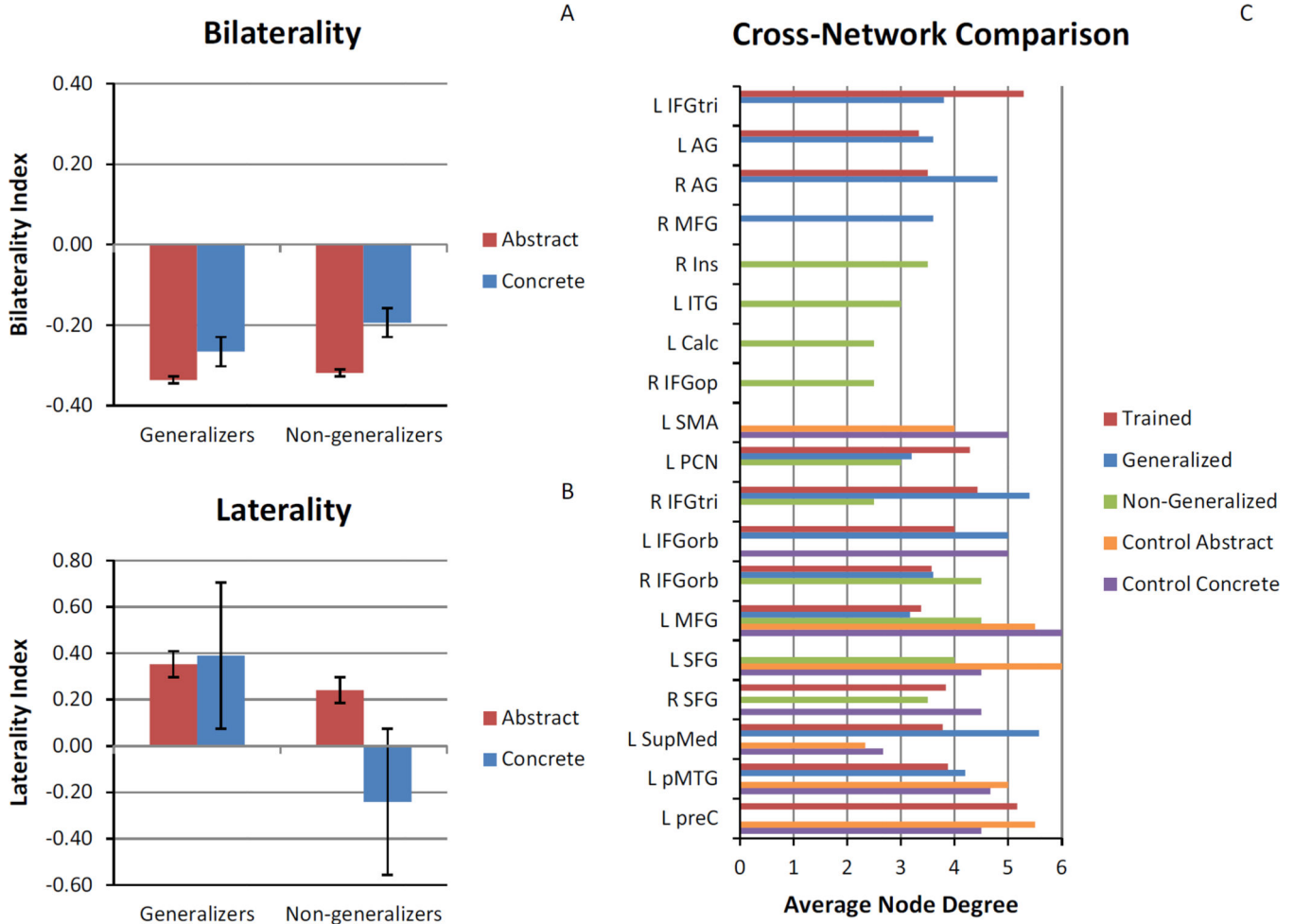
Fig. 6.
Laterality of changes and cross-network comparison of node degree. Panel (A) illustrates the comparison of the bilaterality index between generalizers and nongeneralizers for the trained abstract network and the (non)generalized concrete network. Panel (B) illustrates the comparison of the laterality index between generalizers and non-generalizers for the trained abstract network and the (non)generalized concrete network. Panel (C) illustrates the comparison of the average node degree for increases in connectivity for fROIs among change networks for the majority (at least 2/3) of participants.

Table 1

Key to region abbreviations.

Abbreviation	Region
AG	Angular gyrus
CN	Caudate nucleus
Cun	Cuneus
DMPFC	Dorsomedial prefrontal cortex
FFG	Fusiform gyrus
IFGop	Inferior frontal gyrus pars opercularis
IFGorb	Inferior frontal gyrus pars orbitalis
IFGtri	Inferior frontal gyrus pars triangularis
Ins	Insula
IPL	Inferior parietal lobule
ITG	Inferior temporal gyrus
MCC	Middle cingulate cortex
MFG	Middle frontal gyrus
MOG	Middle occipital gyrus
aMTG	Anterior middle temporal gyrus
pMTG	Posterior middle temporal gyrus
PCC	Posterior cingulate cortex
PCN	Precuneus
PHG	Parahippocampal gyrus
preC	Precentral gyrus
SFG	Superior frontal gyrus
SMA	Supplementary motor area
SMG	Supramarginal gyrus
SOG	Superior occipital gyrus
aSTG	Anterior superior temporal gyrus
aSTS	Anterior superior temporal sulcus
SupMed	Medial superior frontal gyrus

Table 2

Demographic information for all participants.

Patient	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Age	57	56	59	47	48	74	53	69	56	75
Sex	Female	Male	Male	Male	Male	Female	Male	Male	Female	Male
Months post stroke	38	76	23	42	93	134	117	16	7	11
Lesion volume (cc)	62.55	79.79	123.86	14.24	255.95	101.06	163.12	0.33	3.54	108.64
Western Aphasia Battery										
Aphasia quotient	99.2	77.7	78.6	95.5	72.5	90.8	41.7	97.1	84.7	67.4
Aphasia type	Anomic	Conduction	Anomic	Anomic	Conduction	Anomic	Broca's	Anomic	Anomic	TCM
Boston Naming Test	92%	87%	68%	95%	82%	68%	22%	95%	90%	90%
Psycholinguistic Assessment of Language Processing in Aphasia										
Auditory lexical decision: HI	100%	100%	98%	100%	98%	95%	100%	100%	100%	100%
Auditory lexical decision: LI	100%	98%	98%	98%	93%	95%	93%	100%	95%	98%
Visual lexical decision: HI	100%	100%	100%	97%	100%	100%	100%	100%	100%	100%
Visual lexical decision: LI	97%	100%	93%	93%	100%	100%	93%	97%	100%	100%
Auditory synonym judgment: HI	100%	90%	93%	100%	93%	100%	90%	100%	100%	97%
Auditory synonym judgment: LI	97%	90%	77%	90%	77%	93%	67%	93%	100%	90%
Written synonym judgment: HI	100%	97%	87%	97%	100%	87%	87%	100%	100%	100%
Written synonym judgment: LI	100%	83%	77%	100%	87%	93%	63%	93%	100%	100%
Pyramids and Palm Trees										
Pictures	96%	98%	94%	100%	90%	77%	88%	98%	98%	96%
Written words	98%	96%	94%	98%	96%	94%	85%	100%	96%	69%
Cognitive Linguistic Quick Test										
Composite severity	WNL	WNL	WNL	WNL	Mild	Mild	Mild	Mild	WNL	Mild
Baseline categorical word generation										
Concrete	30%	33%	17%	37%	17%	20%	10%	20%	37%	17%
Abstract	10%	3%	3%	3%	3%	3%	0%	3%	3%	0%

Note. TCM = transcortical motor, HI = high imageability, LI = low imageability, WNL = within normal limits.

Table 3

Significant increased activation for abstract and concrete words from pre- to post-treatment for responders and generalizers.

Abstract (n = 9)			Concrete (n = 7)		
t-value	MNI coordinates (x, y, z)	Structure	t-value	MNI coordinates (x, y, z)	Structure
7.34	−6 −27 48	L MCC	7.37	−57 −33 24	L SMG
6.16	54 0 −21	R MTG	7.34	3 −36 48	R MCC
5.88	−39 −54 21	L AG	7.3	−30 −78 21	L MOG
5.65	48 −51 24	R AG	7.13	45 −3 39	R PreC
5.43	−18 −60 27	L Cun	6.75	66 −27 9	R STG
5.28	−12 −63 30	L PCN	6.31	−21 −48 12	L PCN
5.28	−39 −39 −12	L ITG	5.92	−30 6 12	L Ins
5.00	27 −75 42	R SOG	5.8	−30 −75 42	L IPL
4.85	−3 −12 66	L SMA	5.42	24 12 66	R SFG
4.83	−36 −57 51	L IPL			
4.81	−57 −33 24	L SMG			
4.81	−36 12 24	L IFGtri			
4.73	−18 −21 21	L CN			
4.65	−24 30 36	L MFG			
4.63	18 48 27	R SFG			
4.63	54 −30 45	R SMG			
4.55	−12 27 42	L SFG			

Note. One-sample *t*-tests were used to obtain group results of [post-treatment abstract > pre-treatment abstract] and [post-treatment concrete > pre-treatment concrete] contrasts for responders (*n* = 9) and generalizers (*n* = 7). All results shown are at the uncorrected $p < 0.001$ level. See Table 1 for a key to region abbreviations.